(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 997 530 A1

(12) **EUROPEAN PATENT APPLICATION**
published in accordance with Art. 158(3) EPC

(43) Date of publication:
03.05.2000 Bulletin 2000/18

(21) Application number: 98929687.6

(22) Date of filing: 24.06.1998

(51) Int. Cl.^{7}: **C12N 15/54**, C12N 9/12, C12N 15/31, C07K 14/195, C12P 21/02
// C12N15:54, C12R1:01, C12N9:12, C12R1:19, C12N15:31, C12R1:01, C12P21:02, C12R1:19

(86) International application number:
PCT/JP98/02845

(87) International publication number:
WO 99/00506 (07.01.1999 Gazette 1999/01)

(84) Designated Contracting States:
CH DE ES FR GB IT LI NL SE

(30) Priority: 26.06.1997 JP 18749697
21.11.1997 JP 32069297

(71) Applicant: Takara Shuzo Co, Ltd.
Kyoto-shi, Kyoto 612-8061 (JP)

(72) Inventors:
- UEMORI, Takashi
Otsu-shi, Shiga 520-2141 (JP)
- SATO, Yoshimi
Kurita-gun, Shiga 520-3031 (JP)
- FUJITA, Tomoko
Takatsuki-shi, Osaka 569-1144 (JP)
- MIYAKE, Kazue
Uji-shi, Kyoto 611-0011 (JP)
- MUKAI, Hiroyuki
Moriyama-shi, Shiga 524-0102 (JP)
- ASADA, Kiyozo
Koga-gun, Siga 520-3333 (JP)
- KATO, Ikunoshin
Uji-shi, Kyoto611-0028 (JP)

(74) Representative:
Schüssler, Andrea, Dr.
Kanzlei Huber & Schüssler
Truderinger Strasse 246
81825 München (DE)

(54) DNA POLYMERASE-RELATED FACTORS

(57) The present invention relates to a thermostable DNA polymerase-associated factor capable of enhancing DNA synthesizing-activity of a DNA polymerase; a thermostable DNA polymerase-associated factor possessing an activity of binding to a DNA polymerase and a method for producing the same; a gene encoding the DNA polymerase-associated factor; a method of DNA synthesis by using a DNA polymerase in the presence of the DNA polymerase-associated factor; and a kit comprising the DNA polymerase-associated factor. According to the present invention, there can be provided *in vitro* DNA synthesis and a DNA amplification system which are more excellent than conventional techniques by utilizing the DNA polymerase-associated factor of the present invention.

EP 0 997 530 A1

Description

TECHNICAL FIELD

[0001] The present invention relates a DNA polymerase-associated factor. More specifically, the present invention relates to a DNA polymerase-associated factor which is useful for a reagent for genetic engineering and a method for producing the same, and further a gene encoding thereof, and the like.

BACKGROUND ART

[0002] DNA polymerases are useful enzymes for reagents for genetic engineering, and the DNA polymerases are widely used for nucleotide sequencing of DNA, DNA labeling, site-directed mutagenesis, and the like. Also, thermostable DNA polymerases have recently been remarked with the development of the polymerase chain reaction (PCR) method, and various DNA polymerases suitable for the PCR method have been developed and commercialized.

[0003] Presently known DNA polymerases can be roughly classified into four families according to amino acid sequence homologies, among which family A (pol I type enzymes) and family B (α type enzymes) account for the great majority. Although DNA polymerases belonging to each family generally possess mutually similar biochemical properties, detailed comparison reveals that individual enzymes differ from each other in terms of substrate specificity, incorporation efficiency of a substrate analog, primer extensibility and extension rate, mode of DNA synthesis, association of exonuclease activity, optimum reaction conditions of temperature, pH and the like, and sensitivity to inhibitors. Therefore, those possessing most appropriate properties for the applications have been selected among all available DNA polymerases, and the selected DNA polymerase has been used.

[0004] A hyperthermophilic archaebacterium *Pyrococcus furiosus* has produced a DNA polymerase belonging to α type, and its gene has already been isolated [*Nucleic Acids Research*, **21**, 259-265 (1993)].

[0005] As DNA polymerases, in addition to ones expressing their functions with only one kind of an enzyme protein, such as the pol I type enzyme or the α type enzyme, there have been known oligomer enzymes constituted by a large number of subunit proteins. In addition to the protein serving as a DNA polymerase, there have also been known some cases where protein molecules for regulating their functions coexist.

DISCLOSURE OF INVENTION

[0006] An object of the present invention is to provide a thermostable DNA polymerase-associated factor capable of enhancing DNA synthesizing-activity of a DNA polymerase, and a thermostable DNA polymerase-associated factor possessing an activity of binding to a DNA polymerase.

[0007] Another object of the present invention is to provide a gene for the DNA polymerase-associated factor of the present invention.

[0008] Still another object of the present invention is to provide a method for producing the DNA polymerase-associated factor of the present invention.

[0009] Still another object of the present invention is to provide a method of DNA synthesis by using a DNA polymerase in the presence of the DNA polymerase-associated factor of the present invention.

[0010] Still another object of the present invention is to provide a kit comprising the DNA polymerase-associated factor of the present invention.

[0011] According to the present invention, there can be provided *in vitro* DNA synthesis and a DNA amplification system which are more excellent than conventional techniques by utilizing the DNA polymerase-associated factor of the present invention.

[0012] Recently, a novel DNA polymerase having completely no structural homology to conventionally known DNA polymerases has been found by the present inventors from hyperthermophilic archaebacterium *Pyrococcus furiosus* (WO 97/24444 Pamphlet). In this DNA polymerase, two kinds of novel proteins form a complex and exhibit a DNA polymerase activity. In addition, the enzyme exhibits a potent 3' → 5' exonuclease activity and excellent primer extension activity. For example, when the enzyme is used for PCR, a DNA fragment of the size of about 20 kb can be amplified. In this novel DNA polymerase derived from *Pyrococcus furiosus*, although at least two kinds of proteins are essential constituents in the enzyme activity, it has not been elucidated whether or not a constituent protein of the enzyme beside the above exists, or whether or not a factor having an influence on the activity of the enzyme exists.

[0013] As a result of intensive studies, the present inventors have succeeded in isolating a protein binding to the novel DNA polymerase derived from *Pyrococcus furiosus*. Further, they have found that the production of the protein by genetic engineering is made possible by cloning the gene, and moreover that a DNA synthesizing-activity of a DNA polymerase is enhanced.

[0014] In sum, the present invention relates to:

[1] a thermostable DNA polymerase-associated factor capable of enhancing DNA synthesizing-activity of a DNA polymerase;

[2] the DNA polymerase-associated factor according to item [1] above, further possessing an activity of binding to a DNA polymerase;

[3] the DNA polymerase-associated factor according to item [2] above, which possesses an activity of binding to a DNA polymerase comprising a DNA polymerase-constituting protein having the amino acid sequence as shown in SEQ ID NO: 5 or 6 in Sequence Listing;

[4] the DNA polymerase-associated factor according to any one of items [1] to [3] above, comprising at least one of amino acid sequences selected from the group consisting of SEQ ID NOs: 1, 3, 19, 27, 34, 64, 70 and 80 in Sequence Listing, or an amino acid sequence resulting from substitution, deletion, addition or insertion of one or more amino acids in at least one of the amino acid sequences;

[5] a gene encoding a DNA polymerase-associated factor, wherein the factor comprises at least one of amino acid sequences selected from the group consisting of SEQ ID NOs: 1, 3, 19, 27, 34, 64, 70 and 80 in Sequence Listing, or an amino acid sequence resulting from substitution, deletion, addition or insertion of one or more amino acids in at least one of amino acid sequences, and possesses an activity of enhancing DNA synthesizing-activity of a DNA polymerase;

[6] the gene according to item [5] above, comprising a nucleotide sequence selected from the group consisting of SEQ ID NOs: 2, 4, 18, 26, 33, 63, 69 and 79, or a nucleotide sequence resulting from substitution, deletion, addition or insertion of one or more bases in the nucleotide sequence;

[7] a gene capable of hybridizing to the gene of item [5] or [6] above, and encoding a DNA polymerase-associated factor possessing an activity of enhancing DNA synthesizing-activity of a DNA polymerase;

[8] a method for producing a DNA polymerase-associated factor, characterized in that the method comprises culturing a transformant harboring the gene of any one of items [5] to [7] above, and collecting a thermostable DNA polymerase-associated factor capable of enhancing DNA synthesizing-activity of a DNA polymerase from the cultured medium;

[9] a method of DNA synthesis by using a DNA polymerase, characterized in that DNA is synthesized in the presence of the DNA polymerase-associated factor of any one of items [1] to [4] above;

[10] the method of DNA synthesis according to item [9] above, wherein DNA is synthesized in the presence of two or more kinds of DNA polymerase-associated factors;

[11] the method of DNA synthesis according to item [10] above, wherein DNA is synthesized in the presence of F7, PFU-RFC and PFU-RFCLS as a DNA polymerase-associated factor;

[12] the method of DNA synthesis according to any one of items [9] to [11] above, wherein the DNA polymerase is a thermostable DNA polymerase;

[13] the method of DNA synthesis according to item [12] above, wherein the synthesis is carried out by PCR method;

[14] a kit usable for *in vitro* DNA synthesis, comprising the DNA polymerase-associated factor of any one of items [1] to [4] above and a DNA polymerase;

[15] the kit according to item [14] above, further comprising a reagent required for DNA synthesis;

[16] the kit according to item [14] or [15] above, comprising two or more kinds of DNA polymerase-associated factors;

[17] the kit according to item [16] above, comprising F7, PFU-RFC and PFU-RFCLS as a DNA polymerase-associated factor; and

[18] the kit according to any one of items [14] to [17] above, comprising a thermostable DNA polymerase as a DNA polymerase.

BRIEF DESCRIPTION OF DRAWINGS

[0015]

Figure 1 is a drawing showing SDS-PAGE of 7 kinds of proteins (F1, F2, F3, F4, F5, F6 and F7) isolated by an anti-Pfu polymerase C antibody column. The molecular weights on SDS-PAGE are about 55 kDa, about 24 kDa, about 37 kDa, about 19.5 kDa, about 27 kDa, about 64 kDa and about 33 kDa, in a sequential order of F1 to F7.

Figure 2 is a restriction endonuclease map of a DNA insert of the plasmid pF1-4-10 carrying a gene encoding the F1 protein.

Figure 3 is a graph showing a 5' → 3' exonuclease activity of the F1 protein.

Figure 4 is a graph showing a 3' → 5' exonuclease activity of the F1 protein.

Figure 5 is a restriction endonuclease map of a DNA insert of the plasmid pF2172Nh carrying a gene encoding the F2 protein.

Figure 6 is a restriction endonuclease map of a DNA insert of the plasmid pF7-1-8 carrying a gene encoding the F7 protein.

Figure 7 is an autoradiogram showing a primer extension activity of the DNA polymerase when the F7 protein is added.

Figure 8 is an autoradiogram showing a primer extension activity for the higher molecular primer extension reaction product of the DNA polymerase, when the F7 protein is added.

Figure 9 is a restriction endonuclease map of a DNA insert of the plasmid pRFS254NdB carrying a gene encoding the PFU-RFC protein.

Figure 10 shows the analytical results of SDS-PAGE of the protein (F7) isolated by an anti-Pfu DNA polymerase antibody column. The molecular weight of F7 on SDS-PAGE is deduced to be about 33 kDa.

Figure 11 shows the analytical results of DNA polymerase activity of the eluate obtained by subjecting to gel filtration Pfu DNA polymerase and a mixture of Pfu DNA polymerase and F7.

Figure 12 is a restriction endonuclease map of a DNA insert of the plasmid pRFLSNh carrying a gene encoding the PFU-RFCLS protein.

Figure 13 is a restriction endonuclease map around the gene encoding the F5 protein on genomic DNA of *Pyrococcus furiosus*.

Figure 14 shows analytical results of SDS-PAGE of 3 kinds of proteins (PFU-RFCLS, PFU-RFC, F7) isolated by an anti-PFU-RFC antibody column.

Figure 15 is a graph showing DNA polymerase activity when F7 or RFC-N complex is added.

Figure 16 is a restriction endonuclease map of a DNA insert of the plasmid pRFC10 carrying genes encoding PFU-RFCLS and PFU-RFC.

Figure 17 is a graph showing DNA polymerase activity, when F7, or F7 and rRFC-M complex are added.

BEST MODE FOR CARRYING OUT THE INVENTION

1. DNA Polymerase-Associated Factor of the Present Invention

[0016] In the present specification, the term "DNA polymerase-associated factor" means a factor which has effects on a function of a DNA polymerase by coexisting with the DNA polymerase. Concretely, the DNA polymerase-associated factors include a factor possessing an action of enhancing the DNA synthesizing-activity of a DNA polymerase, a factor possessing an activity of binding to a DNA polymerase, and further one possessing both activities, and the like. In addition, the DNA polymerase-associated factor of the present invention is a thermostable protein, which is, for instance, stable against heat treatment at 80°C for 15 minutes. Therefore, the factor can be used for DNA synthesizing-reaction under high-temperature conditions using a thermostable DNA polymerase.

(a) DNA Polymerase-Associated Factor Capable of Enhancing DNA Synthesizing-Activity of DNA Polymerase

[0017] The DNA polymerase-associated factor capable of enhancing DNA synthesizing-activity of a DNA polymerase is not particularly limited, as long as the factor is capable of enhancing DNA synthesizing-activity of a DNA polymerase. For instance, the factor includes proteins comprising an entire or partial sequence of amino acid sequence as shown in at least one sequence selected from the group consisting of SEQ ID NOs: 1, 3, 19, 27, 34, 64, 70 and 80 in Sequence Listing; or functional equivalents thereof comprising an amino acid sequence resulting from substitution, deletion, addition or insertion of one or more amino acids in at least one of the amino acid sequences, and the equivalent possessing an activity of enhancing DNA synthesizing-activity of a DNA polymerase. In the present specification, the term "one or more" refers to a number of one or several or more. In addition, the term "functional equivalent" refers to ones which are substantially equivalent in their functions and activities even though they are structurally different, and the functional equivalents are also encompassed in the DNA polymerase-associated factor of the present invention.

[0018] The DNA polymerase of which activity is enhanced by the DNA polymerase-associated factor of the present invention is not particularly limited. Examples thereof include thermostable DNA polymerases, in particular DNA polymerases derived from hyperthermophilic archaebacterium. Concretely, there can be cited DNA polymerases derived from *Pyrococcus furiosus* (Pfu polymerase C, and the like mentioned below). As described below, the Pfu polymerase C is an enzyme comprising a DNA polymerase-constituting protein having the amino acid sequences as shown in SEQ ID NO: 5 and SEQ ID NO: 6 in Sequence Listing.

[0019] In addition, the DNA polymerase-associated factor of the present invention may be one enhancing only an activity of a particular DNA polymerase, and it is preferably one enhancing its activities against a plural kinds of DNA polymerase from different origins.

[0020] The method for determination of an activity of enhancing DNA synthesizing-activity of a DNA polymerase is not particularly limited, as long as it is one usually employed in the determination of DNA synthesizing-activity of a DNA

polymerase. The activity of enhancing DNA synthesizing-activity can be, for instance, determined by adding the factor when measuring an incorporation activity of the labeled nucleotide into a novel synthesized DNA strand; and comparing the incorporation activity with an activity when the factor is not added. In addition, there can be cited a method for confirmation from the chain length of a novel synthetic DNA strand per unit time or from the amount of PCR amplified product per unit time. As the method for determination of the DNA synthesizing-activity, there can be cited a method described in *DNA Polymerase from Escherichia coli*, published by Harpar and Row, edited by D.R. Davis, 263-276 (authored by C.C. Richardson), and the like.

[0021] Further, in the DNA polymerase-associated factor of the present invention, by a combination of a plurality of the DNA polymerase-associated factors, there can be exhibited an even higher DNA polymerase activity in the coexistent DNA polymerases when compared with that of the single use.

(b) DNA Polymerase-Associated Factor Possessing Activity of Binding to DNA Polymerase

[0022] The DNA polymerase-associated factor possessing an activity of binding to a DNA polymerase is not particularly limited, as long as it possesses an activity of binding to a DNA polymerase. Incidentally, the DNA polymerase-associated factor possessing an activity of binding to a DNA polymerase in the present specification encompasses other substances, for instance, ones having an activity of indirectly binding to a DNA polymerase via other DNA polymerase-associated factors, as well as ones having an activity of directly binding to a DNA polymerase. Examples thereof include proteins comprising an entire or partial sequence of amino acid sequence as shown in at least one sequence selected from the group consisting of SEQ ID NOs: 1, 3, 19, 27, 34, 64, 70 and 80 in Sequence Listing; or functional equivalents thereof comprising an amino acid sequence resulting from substitution, deletion, addition or insertion of one or more amino acids in at least one of the amino acid sequences, and the equivalent possessing an activity of binding to a DNA polymerase. In the present specification, the term "one or more" refers to a number of one or several or more.

[0023] The DNA polymerase binding to the DNA polymerase-associated factor of the present invention, which is not particularly limited, includes, for instance, a thermostable DNA polymerase, in particular DNA polymerases derived from hyperthermophilic archaebacterium. Concretely, there can be cited DNA polymerases derived from *Pyrococcus furiosus* (Pfu polymerase C, and the like). One or both of the DNA polymerase-constituting proteins having the amino acid sequences as shown in SEQ ID NO: 5 and SEQ ID NO: 6 in Sequence Listing are bound to Pfu polymerase C.

[0024] In addition, the DNA polymerase-associated factor of the present invention may be one binding to a particular DNA polymerase, and it is preferably one binding to a plural kinds of DNA polymerase from different origins.

[0025] The method for determination of the binding to a DNA polymerase includes a method comprising mixing the factor with a DNA polymerase, and examining a change in the molecular weight by native gel electrophoresis, gel filtration, and the like; a method for examining the adsorption of the factor to a carrier immobilized to a DNA polymerase, and the like.

[0026] In addition, the DNA polymerase-associated factor comprising the amino acid sequence as shown in SEQ ID NO: 19 in Sequence Listing possesses an exonuclease activity. Therefore, it is considered that the DNA polymerase-associated factor comprising the amino acid sequence as shown in SEQ ID NO: 19 is a protein having a function associated with the action of a DNA polymerase in DNA replication, DNA repair, and the like. Further, as the functional equivalents of the DNA polymerase-associated factor, proteins comprising a partial sequence of the amino acid sequence as shown in SEQ ID NO: 19 in Sequence Listing, or an amino acid sequence resulting from substitution, deletion, addition or insertion of one or more amino acids in at least one of the sequences, wherein the proteins possess an activity of binding to a DNA polymerase, and further similarly possess an exonuclease activity are encompassed in the present invention as the DNA polymerase-associated factor. In the present specification, the term "one or more" refers to a number of one or several or more.

[0027] Incidentally, in the explanation of the DNA polymerase-associated factor of the present invention, the factor is identified as a protein comprising an entire or partial sequence of each of the amino acid sequences as shown in particular SEQ ID NO in Sequence Listing, and the term "protein comprising" as used herein encompasses proteins as described below, which are also encompassed in the present invention. Namely, when a protein is produced by genetic engineering techniques, it is often expressed as a fusion protein. For instance, in order to increase an expression level of the desired protein, the protein is expressed by adding a N-terminal peptide chain derived from other proteins to the N-terminus, or expressed by adding an appropriate peptide chain at N-terminus or C-terminus of the desired protein, and a carrier having affinity with each of the peptide chain is used, whereby facilitating the purification of the desired protein. In the present invention, the fusion proteins mentioned above are also encompassed.

2. Genes Encoding DNA Polymerase-Associated Factor of the Present Invention

(a) Properties of Genes Encoding DNA Polymerase-Associated Factor of the Present Invention

[0028] The genes encoding the DNA polymerase-associated factor of the present invention are those encoding the DNA polymerase-associated factor of the present invention mentioned above, which refers to DNA or RNA. Concretely, the gene includes a gene encoding a DNA polymerase-associated factor, wherein the factor comprises an entire or partial sequence of amino acid sequence as shown in at least one sequence selected from the group consisting of SEQ ID NOs: 1, 3, 19, 27, 34, 64, 70 and 80 in Sequence Listing, or an amino acid sequence resulting from substitution, deletion, addition or insertion of one or more amino acids in at least one of these sequences, and the factor possesses an activity of enhancing DNA synthesizing-activity of a DNA polymerase, or an activity of binding to a DNA polymerase. Concrete examples of such genes include genes encoding a DNA polymerase-associated factor, comprising an entire or partial sequence of nucleotide sequence as shown in at least one sequence selected from the group consisting of SEQ ID NOs: 2, 4, 18, 26, 33, 63, 69 and 79, or a nucleotide sequence resulting from substitution, deletion, addition or insertion of one or more bases in these sequences, wherein the factor possesses an activity of enhancing DNA synthesizing-activity of a DNA polymerase, or an activity of binding to a DNA polymerase. In the present specification, the term "one or more" refers to a number of one or several or more. In the present invention, there can be further cited a gene capable of hybridizing to a DNA of the gene of the present invention, and possessing an activity of enhancing DNA synthesizing-activity, or an activity of binding to a DNA polymerase.

[0029] The term "gene capable of hybridizing (to a gene)" described in the present specification refers to a gene comprising a DNA capable of hybridizing to a DNA of a gene, which is a gene having a nucleotide sequence resembling to the gene. With regard to the gene having a nucleotide sequence resembling to a gene, there is a high possibility of having resemblance to an amino acid sequence of a protein encoded thereby, and additionally having resemblance to a function of the protein. The homology of the nucleotide sequence of the gene can be examined by whether or not a hybrid is formed (the genes being hybridized) with DNAs of both genes or a partial portion thereof under stringent conditions. By utilizing hybridization, a gene encoding a protein having similar functions to a protein encoding the gene can be obtained. In other words, the other genes of the present invention having homologous nucleotide sequences to a gene of the present invention can be obtained by carrying out hybridization by a known method using a DNA of the gene obtained in the present invention, or a partial portion thereof, as a probe. The hybridization can be carried out, for instance, by a method described in *Molecular Cloning: A Laboratory Manual*, 2nd Ed., published by Cold Spring Harbor Laboratory in 1989, edited by T. Maniatis et al., or the like.

[0030] Here, the term "the stringent conditions" refers to conditions in which non-specific hybridization does not take place. Concretely, for instance, there are the following conditions. In other words, a DNA-immobilized membrane is incubated at 50°C for 12 to 20 hours together with a labeled DNA probe in 6 × SSC (wherein 1 × SSC shows 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0) containing 0.5% SDS, 0.1% bovine serum albumin (BSA), 0.1% polyvinyl pyrrolidone, 0.1% Ficol 400, and 0.01% denatured salmon sperm DNA. After termination of the incubation, the membrane is washed, initiating under the conditions of 37°C in 2 × SSC containing 0.5% SDS, the SSC concentration being made variable up to a range of 0.1 × SDS, and the temperature being variable up to a range of 50°C, until a signal ascribed to an immobilized labeled DNA probe can be distinguished from the background.

[0031] In addition, instead of hybridization, there can be utilized a method for gene amplification using a partial sequence of the nucleotide sequence of the gene of the present invention as a primer. For instance, PCR method can be utilized. The PCR conditions can be appropriately set by sequences of primer DNAs or a template DNA. Whether or not the gene obtained as described above encodes a protein having the desired function can be examined by confirming the activity of the resulting protein by expressing a protein encoded by the gene using an appropriate host and an expression system.

[0032] In addition, the method for artificially preparing an amino acid sequence or nucleotide sequence having substitution, deletion, addition, or insertion of one or more in the amino acid sequence or nucleotide sequence in the present invention includes various genetic engineering manipulations described in *Molecular Cloning: A Laboratory Manual*, 2nd Ed., published by Cold Spring Harbor Laboratory in 1989, edited by T. Maniatis et al., or the like. Concrete examples thereof include genetic engineering techniques such as methods for site-directed mutagenesis and cassette mutation methods. By the method for site-directed mutagenesis, an amino acid sequence or nucleotide sequence having one or more substitution, deletion, addition or insertion can be prepared. By the cassette mutation method, there can be prepared an amino acid sequence or nucleotide sequence having a larger region of deletion, addition or insertion as compared with the sequence obtained by the method for site-directed mutagenesis. These modified products described above are also encompassed in the present invention as long as they are functionally equivalent. Further, in the production of a protein by genetic engineering techniques, in a case where a codon used on a naturally occurring gene encoding the desired protein is used at a low frequency, the expression level of the protein may be low. In such a case, the codon is artificially converted to one frequently used in the host without changing the encoded amino acid

sequence, whereby the desired protein is highly expressed (for instance, Japanese Examined Patent Publication No. Hei 7-102146).

(b) Cloning of Gene Encoding DNA Polymerase-Associated Factor of the Present Invention

[0033] Detailed descriptions on the analysis of the resulting clones, the physicochemical properties of the expression product DNA polymerase-associated factor, the elucidation of the functions, and the like will be given hereinbelow.

[0034] As described above, the DNA polymerase-associated factor of the present invention possesses an action of enhancing DNA synthesizing-activity of a DNA polymerase, or a characteristic of binding the factor to a DNA polymerase. Therefore, the factor can be obtained by using these actions as indices.

[0035] The DNA polymerase utilizable in the obtainment of the DNA polymerase-associated factor of the present invention is not particularly limited, and an example thereof includes a *Pyrococcus furiosus*-producing DNA polymerase. As the *Pyrococcus furiosus*-producing DNA polymerase, for instance, there can be used an enzyme comprising a DNA polymerase-constituting protein comprising the amino acid sequence as shown in SEQ ID NO: 5 and/or SEQ ID NO: 6 in Sequence Listing, derived from *Pyrococcus furiosus* DSM3638.

[0036] Incidentally, in the present specification, this enzyme is described as Pfu polymerase C, in order to distinguish with a type DNA polymerase [Pfu DNA polymerase, *Nucleic Acids Research*, **21**, 259-265 (1993)], which has been also found from *Pyrococcus furiosus*. The gene encoding the enzyme is carried by plasmid pFU1001. In addition, a transformant, *Escherichia coli* JM109 transformed with the plasmid, is named and identified as *Escherichia coli* JM109/pFU1001, and deposited under the accession number of FERM BP-5579 with the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, Ministry of International Trade and Industry, of which the address is 1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken (Zipcode 305-8566), Japan, since August 11, 1995 (date of original deposit) under the Budapest Treaty. Therefore, Pfu polymerase C can be obtained by culturing the transformant and purifying from the resulting cultured medium. Incidentally, Pfu polymerase C is an enzyme comprising a DNA polymerase-constituting protein having the amino acid sequence as shown in SEQ ID NO: 5 and/or SEQ ID NO: 6 in Sequence Listing.

[0037] Pfu polymerase C is an enzyme possessing the following properties:

(A) exhibiting a higher activity when the polymerase activity is determined by using as a substrate a complex resulting from annealing of a primer to a single stranded template DNA, as compared to the case where an activated DNA is used as a substrate;

(B) possessing a 3'→5' exonuclease activity;

(C) being capable of amplifying a DNA fragment of about 20 kbp without adding other enzymes, in the case where polymerase chain reaction (PCR) is carried out with λ-DNA as a template under the following conditions: PCR conditions:

a) a composition of reaction mixture: comprising 10 mM Tris-HCl (pH 9.2), 3.5 mM $MgCl_2$, 75 mM KCl, 400 μM each of dATP, dCTP, dGTP and dTTP, 0.01% bovine serum albumin, 0.1% Triton X-100, 5.0 ng/50 μl λ-DNA, 10 pmole/50 μl primer λ1 (SEQ ID NO: 58 in Sequence Listing), primer λ11 (SEQ ID NO: 59 in Sequence Listing), and 3.7 units/50 μl DNA polymerase;

b) reaction conditions: carrying out PCR for 30 cycles, wherein one cycle is 98°C, 10 seconds - 68°C, 10 minutes; and

(D) comprising two kinds of DNA polymerase-constituting proteins corresponding to about 90,000 daltons and about 140,000 daltons on SDS-PAGE, respectively.

[0038] The method of obtaining the DNA polymerase-associated factor of the present invention is not particularly limited. For instance, the factor can be obtained by immobilizing a DNA polymerase, such as Pfu polymerase C, to an appropriate carrier, mixing the DNA polymerase-immobilized carrier with a sample containing the DNA polymerase-associated factor, removing the factor unbound to the carrier, and thereafter eluting the bound carrier. The immobilization of the DNA polymerase to the carrier can be carried out by a known method. Alternatively, an antibody against the DNA polymerase is prepared, and a DNA polymerase may be immobilized by utilizing the antibody-immobilized carrier. For instance, when an anti-Pfu polymerase C antibody is prepared, and the DNA polymerase-associated factor of the present invention is obtained by using the antibody from a sample derived from *Pyrococcus furiosus*, including, for instance, a cell disrupted solution of *Pyrococcus furiosus*, Pfu polymerase C in the sample binds to this antibody when the antibody-immobilized carrier as described above is used. Therefore, it is not necessary to add Pfu polymerase C aside from the sample, so that the DNA polymerase-associated factor can be readily purified.

[0039] The sample used in the obtainment of the DNA polymerase-associated factor of the present invention is not

particularly limited. For instance, there can be used samples derived from microorganisms. Concretely, samples derived from *Pyrococcus furiosus* DSM 3638 can be used. The above strain can be made available from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH. In the case of a cell disrupted solution obtained by culturing the above strain in an appropriate growth medium and preparing from the resulting cultured medium is applied to a column packed with a carrier immobilized with an anti-Pfu polymerase C antibody, several kinds of proteins other than Pfu polymerase C are adsorbed to the column. The gene encoding these proteins can be cloned by the procedures exemplified below.

[0040] First, the above proteins are isolated by a known method, and their N-terminal amino acid sequences are determined. In reference to the amino acid sequences, synthetic oligonucleotides to be used as primers or probes are prepared. Next, PCR is carried out with a genomic DNA of *Pyrococcus furiosus* as a template using this synthetic oligonucleotide as a primer, whereby a DNA fragment carrying the desired gene can be obtained. The conditions for PCR may be appropriately set. Alternatively, a DNA fragment carrying the desired gene can be obtained from a genomic DNA of *Pyrococcus furiosus* by carrying out hybridization using the above oligonucleotide as a probe. In this case, as the hybridization, there can be employed Southern hybridization using a genomic DNA of *Pyrococcus furiosus* obtained by digesting with an appropriate restriction enzyme, colony hybridization using a gene library of a genomic DNA of *Pyrococcus furiosus*, plaque hybridization, dot hybridization, and the like.

[0041] When the DNA fragment as obtained above does not carry a full length of the desired gene, new primers are prepared in reference to the nucleotide sequence of the resulting DNA fragment, and PCR is further carried out, or hybridization is carried out using the resulting DNA fragment or its partial fragment as a probe, whereby a full length of the desired gene can be obtained.

[0042] The manipulations for the PCR and hybridization are not particularly limited, and for instance, they can be carried out in reference to *Molecular Cloning*: A *Laboratory Manual*, 2nd Ed., published by Cold Spring Harbor Laboratory in 1989, edited by T. Maniatis et al.

[0043] When the cell disrupted solution of the strain *Pyrococcus furiosus* DSM 3638 is mixed with the above carrier immobilized with the anti-Pfu polymerase C antibody, there are seven kinds of proteins adsorbed to the carrier as well as Pfu polymerase C. With respect to six kinds among them, in the present invention, their genes have been isolated by the above described manipulations. These proteins are named F1, F2, F3, F4, F5 and F7, respectively, which are the concrete examples of the DNA polymerase-associated factor of the present invention. The nucleotide sequences of an open reading frame of the gene encoding these proteins are shown in SEQ ID NOs: 18, 26, 79, 33, 69 and 2, respectively, in Sequence Listing. In addition, the amino acid sequences of each protein deduced from these nucleotide sequences are shown in SEQ ID NOs: 19, 27, 80, 34, 70 and 1, respectively, in Sequence Listing.

[0044] The cloned gene is introduced into an appropriate host, for instance, *Escherichia coli*, whereby allowing to express a protein encoded thereby. For instance, a transformant of *Escherichia coli* JM109, into which a gene encoding F7 mentioned above is introduced, is named and identified as *Escherichia coli* JM109/pF7-HH-18, and deposited under the accession number of FERM BP-6338 with the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, Ministry of International Trade and Industry, of which the address is 1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken (Zipcode 305-8566), Japan, since June 3, 1997 (date of original deposit) under the Budapest Treaty. F7 can be obtained by culturing the transformant, and recovering a desired product from the resulting culture. It is elucidated in the present invention that the F7 as obtained above enhances activities of a type polymerase (Pfu DNA polymerase) derived from *Pyrococcus furiosus* and two kinds of DNA polymerases [*J. Bacteriol.*, **177**, 2164-2177 (1995)] derived from *Pyrodictium occultum*, in addition to Pfu polymerase C used in protein isolation.

[0045] In addition, there are also elucidated that each of F1, F2, F3, F4 and F5 mentioned above enhances an activity of Pfu polymerase C and Pfu DNA polymerase.

[0046] When the amino acid sequence of the protein derived from the above strain *Pyrococcus furiosus* DSM 3638 is compared with an amino acid sequence of a known protein, F1 has homologies to a single-stranded DNA-specific exonuclease derived from *Haemophilis influenzae* [Science, **269**, 496-512 (1995)]. F3 has homologies to *Mycoplana ramosa*-derived acetylpolyamine aminohydrase [*Journal of Bacteriology*, **178**, 5781-5786 (1996)] and human histone deacetylase [Science, **272**, 408-411 (1996)]. In addition, F7 has homologies to the proliferating cell nuclear antigen (PCNA) involved in the DNA replication in eukaryotes [*EMBO J.*, **11**, 5111-5120 (1995); *Nucleic Acids Research*, **18**, 1363-1381 (1990); *Proc. Natl. Acad. Sci. USA*, **84**, 1575-1579 (1987)]. F2, F4 and F5 have not been found to have homologies to a known protein.

[0047] There has been reported that PCNA forms a complex with a replication factor C (RFC, RF-C) to be involved in DNA synthesis [*Journal of Biochemistry*, **68**, 1542-1548 (1996)]. Therefore, even in *Pyrococcus furiosus*, it is expected that a protein corresponding to RFC is expressed, and that the protein is involved in DNA synthesis reaction together with F7 mentioned above. A further excellent effect of enhancing DNA polymerase synthesizing-activity can be obtained by collecting this protein, and for instance, adding the resulting protein together with F7 mentioned above in the reaction system for DNA polymerase. The gene encoding an RFC homolog of *Pyrococcus furiosus* can be obtained by the steps described below.

[0048] An entire nucleotide sequence of chromosomal DNA of archaebacteria *Methanococcus jannaschii* has been already elucidated [Science, **273**, 1058-1073 (1996)], and the nucleotide sequences carry the gene encoding a protein which is considered to be a homolog of PCNA and RFC. The amino acid sequence encoded by the gene of a homolog of RFC small subunit and large subunit of the strain is compared with the amino acid sequence encoded by a known RFC small subunit gene [*Nucleic Acids Research*, **21**, 1-3 (1993); *Nucleic Acids Research*, **22**, 1527-1535 (1994)], thereby examining for the amino acid sequences of high homologies. A synthetic oligonucleotide can be prepared in reference to the above, the oligonucleotide usable as a primer or probe for obtaining a gene fragment encoding RFC small subunit and large subunit. Subsequently, by the manipulations employed for the obtainment of the gene encoding any one of F1 to F7 mentioned above using the oligonucleotide, there can be obtained, for instance, a gene encoding PFU-RFC, which is a homolog of RFC small subunit, and a gene encoding PFU-RFCLS, which is a homolog of RFC large subunit, each derived from *Pyrococcus furiosus*.

[0049] The nucleotide sequence of the gene encoding the PFU-RFC obtained as above is determined, and an amino acid sequence deduced to be encoded thereby is examined, and the amino acid sequence is compared with the amino acid sequence of a known RFC small subunit. As a result, there has been elucidated that an intervening sequence (intein) is present in the amino acid sequence.

[0050] A region corresponding to intein is eliminated from the gene, whereby a gene comprising PFU-RFC in an expressible state can be obtained. The nucleotide sequence of an open reading frame of a region encoding PFU-RFC in the gene and the amino acid sequence of PFU-RFC deduced from the nucleotide sequence are shown in SEQ ID NOs: 4 and 3, respectively, in Sequence Listing. In addition, the nucleotide sequence of an open reading frame encoding PFU-RFCLS in the PFU-RFCLS gene and the amino acid sequence of the protein encoded thereby are shown in SEQ ID NOs: 63 and 64, respectively, in Sequence Listing. Both of these proteins are also one of concrete examples of the DNA polymerase-associated factor of the present invention.

[0051] Further, a plasmid to be used for expression of PFU-RFC can be constructed by using the gene. Such an expression plasmid includes plasmid pRFS254SNc. In addition, a transformant of *Escherichia coli* JM109, into which the plasmid is introduced, is named and identified as *Escherichia coli* JM109/pRFS254SNc, and deposited under the accession number of FERM BP-6339 with the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, Ministry of International Trade and Industry, of which the address is 1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken (Zipcode 305-8566), Japan, since June 3, 1997 (date of original deposit) under the Budapest Treaty. PFU-RFC can be obtained by culturing the transformant, and collecting from the resulting culture. With regard to PFU-RFC, it is observed that the PFU-RFC enhances an activity of a DNA polymerase when used alone, and that the PFU-RFC exhibits synergistic effects in enhancing actions as compared to a case where each protein is added alone when used in combination of F7 above.

[0052] In addition, a transformant resulting from introduction of both PFU-RFC gene and PFU-RFCLS gene is prepared, whereby a complex formed with PFU-RFC and PFU-RFCLS (hereinafter referred to as "holo-RFC"; in particular, holo-RFC produced by genetic engineering is referred to as "rRFC-M complex") can be expressed. The complex is capable of enhancing an activity of a DNA polymerase, which particularly shows high effects when used in combination with F7 mentioned above.

[0053] The above PFU-RFC and PFU-RFCLS can be further allowed to enhance a DNA polymerase activity by using a mixture with F7. In this case, a mixture of the holo-RFC (or rRFC-M complex) with F7 may be used, or a complex formed by PFU-RFC, PFU-RFCLS and F7 (RFC-N complex) may be used.

[0054] As explained above, the present invention provides a DNA polymerase-associated factor capable of enhancing DNA synthesizing-activity of a DNA polymerase, and a gene encoding the factor. The factor can be produced by genetic engineering by utilizing the gene. Further, a gene encoding a protein having an equivalent function with the DNA polymerase-associated factor of the present invention can be also obtained by genetic engineering techniques by utilizing the gene.

[0055] The DNA polymerase-associated factor of the present invention comprises a known protein involved in the DNA synthesis reaction as described above. Examples of such known proteins include ones homologous to proteins such as PCNA and RFC derived from eukaryotes. It has been said that these proteins such as PCNA and RFC form a complex to be involved in the DNA synthesis reaction with DNA polymerase δ [*Journal of Biochemistry*, **68**, 1542-1548 (1996)]. However, the DNA polymerase-associated factor disclosed in the present invention is capable of enhancing an activity of a DNA polymerase with not only the complex, but also individual factors alone. Also, the factor exhibits an effect on a DNA polymerase which is structurally different from DNA polymerase δ.

[0056] The present invention can be utilized in various processes utilizing a DNA polymerase, including, for instance, nucleotide sequencing for DNA, DNA labeling, DNA amplification by PCR, and the like. The DNA polymerase-associated factor of the present invention is added to a reaction system for a DNA polymerase, whereby particularly showing an improvement in an activity of extension of DNA strand from the primer. In addition, since the factor has a high thermostability, it can be utilized for PCR, particularly for PCR in which an amplification of a long chain DNA is desirable.

[0057] Further, among the DNA polymerase-associated factors of the present invention, ones having an activity of binding to a DNA polymerase can be used for detection, purification, and the like, of the DNA polymerase. For instance, the factor can efficiently purify the bound DNA polymerase by subjecting it to affinity chromatography using a carrier to which the DNA polymerase-associated factor of the present invention is bound.

3. Method for Producing DNA Polymerase-Associated Factor of the Present Invention

[0058] One of the features of the method for producing a DNA polymerase-associated factor of the present invention resides in that the method comprises culturing a transformant harboring the gene of the present invention, and collecting from the cultured medium a thermostable DNA polymerase-associated factor capable of enhancing DNA synthesizing-activity of a DNA polymerase, or possessing an activity of binding to a DNA polymerase.

[0059] In the method for producing a DNA polymerase-associated factor of the present invention, a generally employed method for purification of proteins can be applied. For instance, a DNA encoding the DNA polymerase-associated factor of the present invention is ligated to an expression vector, whereby being overexpressed under the control of a promoter of the expression vector. In addition, the DNA polymerase-associated factor of the present invention can be easily collected from a transformant harboring the gene of the present invention by a process comprising ligating a DNA encoding the DNA polymerase-associated factor of the present invention to a DNA encoding a protein such as glutathione reductase and β-galactosidase or to a DNA encoding histidine tag, to be expressed as a fusion protein. The fusion protein mentioned above can be easily isolated by using usually employed affinity column chromatography, such as nickel column. In the fusion protein mentioned above, the DNA polymerase-associated factor can be separated from a protein such as glutathione reductase or β-galactosidase by a conventional method.

[0060] In addition, the expressed DNA polymerase-associated factor of the present invention can be obtained in the same manner as the method for obtaining the DNA polymerase-associated factor of the present invention from *Pyrococcus furiosus*, the method comprising immobilizing a DNA polymerase, such as Pfu polymerase C, to an appropriate carrier, mixing the DNA polymerase-immobilized carrier with a sample containing the DNA polymerase-associated factor, removing ones unbound to the carrier, and eluting one bound thereto.

4. Method of DNA Synthesis

[0061] One of the great features of the method of DNA synthesis of the present invention resides in that a DNA is synthesized using a DNA polymerase in the presence of the DNA polymerase-associated factor of the present invention mentioned above. In the method of DNA synthesis of the present invention, a DNA is synthesized using a DNA polymerase in the presence of the DNA polymerase-associated factor of the present invention, whereby a long chain DNA of about 20 kb can be amplified.

[0062] The DNA polymerase-associated factor usable in the method of DNA synthesis of the present invention includes F1, F2, F3, F4, F5, F7, PFU-RFC, PFU-RFCLS and the like. In the method of DNA synthesis of the present invention, the DNA polymerase-associated factor may be used alone or in admixture of two or more kinds. In the method of DNA synthesis of the present invention, an even longer DNA fragment can be synthesized as compared with the length of the DNA fragment obtained in the conventional method of DNA synthesis by, for instance, using three kinds of the DNA polymerase-associated factors F7, PFU-RFC and PFU-RFCLS. In the method of DNA synthesis of the present invention, the three kinds of the DNA polymerase-associated factors may be used by mixing the three kinds each supplied singly, or they may be used in admixture two kinds of F7 and holo-RFC constituted by PFU-RFC and PFU-RFCLS (rRFC-M complex). Further, the three kinds of the DNA polymerase-associated factors may be used as a complex constituted by F7, PFU-RFC and PFU-RFCLS (RFC-N complex).

[0063] The DNA polymerase used in the method of DNA synthesis of the present invention includes DNA polymerases such as pol I derived from *E. coli*; and thermostable DNA polymerases such as Tth DNA polymerase derived from *Thermus thermophilus*, Taq DNA polymerase derived from *Thermus aquaticus*, and Pfu DNA polymerase derived from *Pyrococcus furiosus*.

[0064] In addition, in the method of DNA synthesis of the present invention, a DNA can be synthesized by PCR method using the DNA polymerase mentioned above.

[0065] In the method of DNA synthesis of the present invention, the amount of the DNA polymerase-associated factor of the present invention to be present is not particularly limited, and an amount sufficient for exhibiting an activity of enhancing synthesizing-activity of the DNA polymerase may be used.

5. Kit Comprising DNA Polymerase-Associated Factor of the Present Invention

[0066] The DNA polymerase-associated factor of the present invention can be utilized in various reactions in which a DNA polymerase is used. Therefore, the DNA polymerase-associated factor of the present invention is attached to a

kit usable for *in vitro* DNA synthesis, including, for instance, a kit for nucleotide sequencing of DNA by the dideoxy method, a kit for DNA labeling, a PCR kit, whereby improving the performance of each of these kits. Besides ones containing the DNA polymerase and the DNA polymerase-associated factor of the present invention, the kit as described above may comprise a reagent required for the reaction of a DNA polymerase, the reagent including, for instance, dNTP and $MgCl_2$. The DNA polymerase-associated factor contained in the kit of the present invention includes F1, F2, F3, F4, F5, F7, PFU-RFC and PFU-RFCLS. In the kit of the present invention, the DNA polymerase-associated factor may be used alone or in admixture of two or more kinds. It is preferable to use three kinds of the DNA polymerase-associated factors F7, PFU-RFC and PFU-RFCLS. Each of the three kinds of the DNA polymerase-associated factors may be used by mixing each of the three kinds supplied singly. Also, there may be used in admixture of two kinds F7 and holo-RFC constituted by PFU-RFC and PFU-RFCLS (rRFC-M complex). Further, the three kinds of the DNA polymerase-associated factors may be used as a complex constituted by F7, PFU-RFC and PFU-RFCLS (RFC-N complex). The DNA polymerase contained in the kit of the present invention also includes DNA polymerases such as pol I derived from *E. coli*; and thermostable DNA polymerases such as Tth DNA polymerase derived from *Thermus thermophilus*, Tag DNA polymerase derived from *Thermus aquaticus*, Pfu DNA polymerase derived from *Pyrococcus furiosus*. In the kit of the present invention, it is preferable that the kit comprises a thermostable DNA polymerase. The kit of the present invention is used for the method of DNA synthesis, whereby a high molecular DNA can be synthesized more simply.

EXAMPLES

[0067] The present invention is hereinafter described by means of the following examples, but the scope of the present invention is not limited only to those examples.

Example 1

(1) Preparation of *Pyrococcus furiosus* Genomic DNA

[0068] *Pyrococcus furiosus* DSM3638 was cultured in the following manner.

[0069] A medium having a composition comprising 1% trypton, 0.5% yeast extract, 1% soluble starch, 3.5% Jamarin S Solid (manufactured by Jamarin Laboratory), 0.5% Jamarin S Liquid (manufactured by Jamarin Laboratory), 0.003% $MgSO_4$, 0.001% NaCl, 0.0001% $FeSO_4 \cdot 7H_2O$, 0.0001% $CoSO_4$, 0.0001% $CaCl_2 \cdot 7H_2O$, 0.0001% $ZnSO_4$, 0.1 ppm $CuSO_4 \cdot 5H_2O$, 0.1 ppm $KAl(SO_4)_2$, 0.1 ppm H_3BO_3, 0.1 ppm $Na_2MoO_4 \cdot 2H_2O$, and 0.25 ppm $NiCl_2 \cdot 6H_2O$ was placed in a two-liter medium bottle and sterilized at 120°C for 20 minutes. After sparging with nitrogen gas thereinto for removal of dissolved oxygen, the above strain was inoculated into the resulting medium. Thereafter, the medium was cultured by allowing to stand at 95°C for 16 hours. After termination of the cultivation, cells were harvested by centrifugation.

[0070] The harvested cells were then suspended in 4 ml of 0.05 M Tris-HCl (pH 8.0) containing 25% sucrose. To this suspension, 0.8 ml of lysozyme [5 mg/ml, 0.25 M Tris-HCl (pH 8.0)] and 2 ml of 0.2 M EDTA were added, and the resulting mixture was incubated at 20°C for 1 hour. Thereafter, 24 ml of an SET solution [150 mM NaCl, 1 mM EDTA, and 20 mM Tris-HCl (pH 8.0)] was added thereto, and 4 ml of 5% SDS and 400 μl of proteinase K (10 mg/ml) were further added to the resulting mixture. Thereafter, the resulting mixture was reacted at 37°C for 1 hour. After termination of the reaction, phenol-chloroform extraction and subsequent ethanol precipitation were carried out to prepare about 3.2 mg of genomic DNA.

(2) Preparation of Cosmid DNA Library

[0071] Four hundred micrograms of the genomic DNA from *Pyrococcus furiosus* DSM3638 was partially digested with Sau3A1 and fractionated by size into 35 to 50 kb fractions by density gradient ultracentrifugation method. Next, 1 μg of triple helix cosmid vector (manufactured by Stratagene) was digested with *Xba*I, and thereafter dephosphorylated using an alkaline phosphatase (manufactured by Takara Shuzo Co., Ltd.), and further digested with *Bam*HI. The resulting treated vector was mixed with 140 μg of the above 35 to 50 kb DNA fractions, and the mixture was subjected to ligation reaction. The cosmid carrying the genomic DNA fragment from *Pyrococcus furiosus* was packaged into lambda phage particles by *in vitro* packaging method using the resulting reaction mixture and "GIGAPACK GOLD" (manufactured by Stratagene), to prepare cosmid library. Subsequently, a portion of this library was transduced into *E. coli* DH5αMCR (manufactured by BRL). Five hundred clones were selected from the resulting transformants, each named as Cosmid Clone No. 1 to No. 500. Further, a cosmid DNA was prepared from each of these clones. Several of them out of the resulting cosmid DNAs were selected and digested with a restriction enzyme to confirm the presence of an insert of an appropriate size.

(3) Cloning of Pfu Polymerase C Gene

[0072] There was prepared as a reaction solution 20 mM Tris-HCl (pH 7.7), 2 mM $MgCl_2$, 2 mM 2-mercaptoethanol, 0.2 mg/ml activated DNA, 40 μM each of dATP, dCTP, dGTP and dTTP, 60 nM [^{3}H]-dTTP (manufactured by Amersham). To 45 μl of the reaction solution was added a 1 μl extract in 5 clone equivalent (5 μl) derived from each clone of the above cosmid DNA library, and the mixture was reacted at 75°C for 15 minutes. Thereafter, a 40 μl aliquot of this reaction mixture was then spotted onto DE paper and washed with 5% Na_2HPO_4 five times. The remaining radioactivity on the DE paper was determined using a liquid scintillation counter. Primary determination was carried out with one group consisting of 5 clones. The group found to have some activities was subsequently separated into one clone each from the 5 clones, and secondary determination was then carried out. Since it had been already known from a hybridization test with the gene as a probe that those clones in the cosmid DNA library containing a known DNA polymerase gene were Clone Nos. 57, 154, 162 and 363, there were obtained five clones of Clone Nos. 41, 153, 264, 462 and 491 possessing DNA synthesizing-activity other than those clones.

[0073] Cosmids were isolated from the above five clones, and each isolated cosmid was digested with *Bam*HI. When examining the resulting electrophoretic patterns, there were found several mutually common bands, predicting that those five clones recombine regions with overlaps and slight shifts. With this finding in mind, the restriction endonuclease map was prepared for the DNA inserts in Clone Nos. 264 and 491. On the basis of the resulting restriction endonuclease map, various DNA fragments of 10 kbp or so in length were cut out from the cosmid derived from Clone 264 or 491. The fragments were then subcloned into pTV118N or pTV119N vector (manufactured by Takara Shuzo Co., Ltd.). The thermostable DNA polymerase activity was measured for the resulting transformant harboring the recombinant plasmid obtained. As a result, it was found that a gene for producing a highly thermostable DNA polymerase was present on an *Xba*I-*Xba*I fragment of about 10 kbp. A plasmid resulting from incorporation of the XbaI-XbaI fragment into pTV118N vector was then named as plasmid pFU1001, and the *Escherichia coli* JM109 transformed with the plasmid was named as *Escherichia coli* JM109/pFU1001 (FERM BP-5579).

(4) Analysis of DNA Polymerase-Constituting Protein of Pfu Polymerase C

[0074] The above *Xba*I-*Xba*I fragment containing the DNA polymerase gene, was again cut out from the above plasmid pFU1001 with *Xba*I, and blunt-ended using DNA blunting kit (manufactured by Takara Shuzo Co., Ltd.). The resultant was then ligated to new pTV118N vector, previously linearized with *Sma*I, to yield plasmids for preparing deletion mutants. The resulting plasmids were named as pFU1002 and pFU1003, respectively, in accordance with the orientations of the inserts. Deletion mutants were prepared from sequentially deleting from both ends of the DNA insert using these plasmids. Kilo-Sequence Deletion kit (manufactured by Takara Shuzo Co., Ltd.) applying Henikoff's method (*Gene*, 28, 351-359) was used for the above preparation. The 3'-overhanging and 5'-overhanging restriction enzymes used were *Pst*I and *Xba*I, respectively. The nucleotide sequence of the insert was determined by the dideoxy method using BcaBEST dideoxy sequencing kit (manufactured by Takara Shuzo Co., Ltd.) with the various deletion mutants as templates. The resulting nucleotide sequence was analyzed, and as a result, there were found six open reading frames (ORFs). The thermostable DNA polymerase activity was determined using the above various deletion mutants. The results demonstrated that the translation products of the ORF3 and the ORF4 were important in the exhibition of the DNA polymerase activity. The amino acid sequence of the ORF3 is shown in SEQ ID NO: 5 in Sequence Listing, and the amino acid sequence is shown in SEQ ID NO: 6 in Sequence Listing, respectively. In other words, the Pfu polymerase C is an enzyme comprising two kinds of the DNA polymerase-constituting proteins having amino acid sequences as shown in SEQ ID NO: 5 and SEQ ID NO: 6 in Sequence Listing, respectively.

Example 2

(1) Preparation of Pfu Polymerase C

[0075] Pfu polymerase C used as an antigen was prepared in the following manner. *Escherichia coli* JM109/pFU1001 was cultured in 2 liter of LB medium (1.0% trypton, 0.5% yeast extract, 0.5% NaCl, pH 7.2) containing 100 μg/ml ampicillin. When the turbidity of the culture reached 0.6 in A_{600}, an inducer, isopropyl-β-D-thiogalactoside (IPTG) was added so as to have a final concentration of 1 mM, and cultured for additional 16 hours. After harvesting, the harvested cells were suspended in 37 ml of sonication buffer [50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol, 10% glycerol, 2 mM PMSF (phenylmethanesulfonyl fluoride)], and the suspension was treated with an ultrasonic disrupter. The supernatant resulting from centrifugation of the disrupted solution at 12,000 rpm for 10 minutes was heat-treated at 80°C for 15 minutes. Thereafter, centrifugation was again carried out at 12,000 rpm for 10 minutes and the supernatant was recovered, to yield 33 ml of a heat-treated supernatant. Subsequently, the above solution was subjected to 2-hour dialysis for 4 times with 2 liter of buffer A [50 mM potassium phosphate, pH 6.5, 2 mM 2-mercaptoeth-

anol, 10% glycerol] as a dialysate. After dialysis, 32 ml of the enzyme solution was applied to RESOURCE Q column (manufactured by Pharmacia) which was previously equilibrated with buffer A, and the applied solution was chromatographed using FPLC system (manufactured by Pharmacia). The elution was carried out on a linear concentration gradient from 0 to 500 mM NaCl. A fraction having a DNA polymerase activity was eluted at 340 mM NaCl.

[0076] Ten milliliters of an enzyme solution obtained by collecting an active fraction was concentrated by using Centriflow CF-50 (manufactured by Grace Japan), and the concentrated enzyme solution was then subjected to exchange with buffer A containing 150 mM NaCl with PD-10 column (manufactured by Pharmacia) to yield 3.5 ml of an enzyme solution. The resulting enzyme solution was then applied to HiTrap Heparin column (manufactured by Pharmacia), previously equilibrated with the same buffer. An active fraction eluted at a concentration of 400 mM NaCl was obtained by eluting with a linear concentration gradient from 150 to 650 mM NaCl using FPLC system. Five milliliters of this fraction was concentrated by ultrafiltration using Centricon-10 (manufactured by Amicon), and 120 μl of the resulting concentrate was applied to Superose 6 gel filtration column (manufactured by Pharmacia), previously equilibrated with 50 mM potassium phosphate buffer (pH 6.5) containing 75 mM NaCl and 2 mM 2-mercaptoethanol, and the elution was carried out with the same buffer. As a result, a fraction having a DNA polymerase activity was eluted at positions corresponding to retention times of 34.7 minutes and 38.3 minutes. The fraction eluted at the position of 38.3 minutes was concentrated, and the resulting concentrate was used as an antigen in the preparation of an anti-Pfu polymerase C polyclonal antibody.

[0077] Incidentally, in the purification of the above Pfu polymerase C, the enzyme activity was determined in the following manner. An activated calf thymus DNA (manufactured by Worthington) (activated DNA) was used as a substrate. Determinations of DNA activation and DNA polymerase activity were carried out by the method described in *DNA Polymerase from Escherichia coli*, 263-276 (authored by C.C. Richardson), published by Harper & Row, edited by D.R. Davis. To 5 μl of a sample of which the activity was to be determined was added 45 μl of a reaction solution [20 mM Tris-HCl (pH 7.7), 15 mM $MgCl_2$, 2 mM 2-mercaptoethanol, 0.2 mg/ml activated DNA, 40 μM each of dATP, dCTP, dGTP and dTTP, 60 nM [^{3}H]-dTTP (manufactured by Amersham)]. The resulting mixture was reacted at 75°C for 5 minutes. A 40 μl portion of this reaction mixture was then spotted onto DE paper (manufactured by Whatman) and washed with 5% Na_2HPO_4 five times. The remaining radioactivity on the DE paper was determined using a liquid scintillation counter. The amount of enzyme which incorporated 10 nmol of [^{3}H]-dTMP per 30 minutes into the substrate DNA, determined by the above-described enzyme activity determination method, was defined as one unit of the enzyme.

(2) Preparation of Anti-Pfu Polymerase C Antibody

[0078] The above Pfu polymerase C preparation was diluted with 50 mM potassium phosphate, pH 6.5, 2 mM 2-mercaptoethanol, and 75 mM NaCl so as to have a concentration of 1 mg/100 μl. Thereto was added an equal volume of complete Freund's adjuvant, and the mixture was emulsified. The resulting emulsion was subcutaneously injected at 50 μl per injection to rabbit 4 times in 3-week intervals. Whole blood was extracted 10 days after the final immunization, and the extracted blood was allowed to stand at room temperature for 60 minutes. Thereafter, the blood was centrifuged to yield 60 ml of antisera containing anti-Pfu polymerase C polyclonal antibody. To 20 ml of the antisera was added 20 ml of saturated ammonium sulfate solution. The mixture was gently stirred at 4°C for 45 minutes, and centrifuged. The resulting precipitate was suspended in 5 ml of 20 mM sodium phosphate buffer, pH 7.0, and the suspension was subjected to a 2-hour dialysis for 3 times using 2 liters of the same buffer as a dialysate. After dialysis, 14 ml of the solution was applied to protein A column (manufactured by Pharmacia), previously equilibrated with 20 mM sodium phosphate buffer (pH 7.0), washed with the same buffer, and then eluted with 0.1 M sodium citrate buffer (pH 3.0). The eluted anti-Pfu polymerase C polyclonal antibody was neutralized with 1 M Tris-HCl, pH 9.0, and concentrated with Centriflow CF-50, and subjected to exchange with coupling buffer (0.5 M NaCl, 0.2 M $NaHCO_3$, pH 8.3) with PD-10 column (manufactured by Pharmacia), to prepare a solution containing anti-Pfu polymerase C polyclonal antibody.

(3) Preparation of Anti-Pfu Polymerase C Antibody Column

[0079] HiTrap NHS-activated column (manufactured by Pharmacia) was washed with 6 ml of 1 mM HCl, and 0.9 ml of the above anti-Pfu polymerase C polyclonal antibody solution (containing 3.6 mg equivalent of the anti-Pfu polymerase C polyclonal antibody) was then applied to HiTrap NHS-activated column. After allowing to stand at room temperature for 1 hour, the resulting column was washed with 3 ml of the coupling buffer. Subsequently, the column was sequentially washed with 6 ml of blocking buffer (0.5 M Tris-HCl, pH 8.3, 0.5 M NaCl), 6 ml of buffer B (0.1 M sodium acetate, pH 4.0, 0.5 M NaCl), and 6 ml of the blocking buffer, and the resulting mixture was allowed to stand at room temperature for 30 minutes. Further, the column was washed with 6 ml of buffer B, 6 ml of the blocking buffer, and 6 ml of buffer B, and thereafter the column was equilibrated with 50 mM Tris-HCl, pH 8.0, to prepare an anti-Pfu polymerase C antibody column.

Example 3

(1) Purification of Complex Comprising Pfu Polymerase C Using Anti-Pfu Polymerase C Antibody Column

[0080] *Pyrococcus furiosus* DSM3638 was cultured in two medium bottles for 16 hours in the same manner as the method described in Example 1. After harvesting, cells were suspended in 34.7 ml of buffer C (50 mM Tris-HCl, pH 8.0, 1 mM ATP) containing 2 mM PMSF, and the suspension was treated with an ultrasonic disrupter. The disrupted solution was centrifuged at 12,000 rpm for 10 minutes, and 46 ml of the supernatant obtained was applied to an anti-Pfu polymerase C antibody column, previously equilibrated with buffer C. After the column was washed with buffer C, the complex comprising Pfu polymerase C was eluted with elution buffer (0.1 M glycine-HCl, pH 2.5, 1 mM ATP). After neutralization with 1 M Tris-HCl, pH 9.0, the eluate was concentrated using Centriflow CF-50 to yield a Pfu polymerase C complex concentrate.

(2) Analysis of Pfu Polymerase C Complex

[0081] The Pfu polymerase C complex concentrate was subjected to SDS-PAGE (12.5% polyacrylamide gel; 25 mM Tris-HCl, 192 mM glycine, 0.1% SDS, pH 8.4 being used as electrophoresis buffer). The gel obtained was analyzed by Western blotting using the anti-Pfu polymerase C antibody by the method shown below. After SDS-PAGE, the gel was immersed in blotting buffer 1 (25 mM Tris-HCl, 20% methanol, pH 9.4) containing 40 mM ε-amino-n-caproic acid. Next, filter papers immersed in blotting buffer 2 (0.3 M Tris-HCl, 20% methanol, pH 10.4), filter papers immersed in 25 mM Tris-HCl and 20% methanol, pH 10.4, a PVDF membrane immersed in blotting buffer 1 containing 40 mM ε-amino-n-caproic acid, the above gel, and filter papers immersed in blotting buffer 1 containing 40 mM ε-amino-n-caproic acid were overlaid on semi-dry blotting apparatus (manufactured by Scientific), and blotting was carried out at 2 mA/cm^2 for 1 hour. This PVDF membrane was immersed in Block Ace (manufactured by Snow Brand Milk Products Co., Ltd.) containing 0.01% thimerosal, shaken for 10 minutes, and thereafter the membrane was immersed in an anti-Pfu polymerase C antiserum, previously diluted 1,000 fold with Block Ace containing 0.01% thimerosal. After allowing to stand at room temperature for 1 hour, the membrane was washed thrice for 10 minutes with TBS buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl) containing 0.02% Tween-20 and further washed with TBS buffer. The membrane was then immersed in a peroxidase-labeled anti-rabbit IgG (Fc) antibody (manufactured by Organon-Technica), previously diluted 5,000 fold with Block Ace containing 0.01% thimerosal. After allowing to stand at room temperature for 1 hour, the PVDF membrane was washed thrice for 10 minutes with TBS buffer containing 0.02% Tween-20 and further washed with TBS buffer. Thereafter, the membrane was immersed in Konica Immunostain HRP-1000 (manufactured by Konica Corporation) to allow color development. From the results of staining of the gel after SDS-PAGE with Coomassie Brilliant Blue R-250, shown in Figure 1, and the results of the Western blotting mentioned above, it was elucidated that the above complex fraction contained seven kinds of proteins (F1 to F7 in Figure 1) unreactive with the anti-Pfu polymerase C antibody.

[0082] Since the bands unreactive with the anti-Pfu polymerase C antibody are considered to be proteins adsorbed to the column via Pfu polymerase C, N-terminal amino acid sequences of these proteins were analyzed by the method described below. The Pfu polymerase C complex concentrate obtained in Example 3(1) was subjected to SDS-PAGE and blotted onto a PVDF membrane in the same manner as above. After this membrane was stained with Coomassie Brilliant Blue R-250, the desired bands were cut out. The N-terminal amino acid sequences of the desired proteins were determined by automatic Edman decomposition with G1000A Protein Sequencer (manufactured by Hewlett-Packard Company) using these membrane fragments as samples. The results are shown in Table 1. The N-terminal amino acid sequences obtained, F1 to F5 and F7, are shown in SEQ ID NOs: 7 to 12, respectively, in Sequence Listing.

Table 1

Sample	N-Terminal Amino Acid Sequence
F1	MDKEGFLNKVREAVDVVKLH
F2	MFTGKVLIPVKVLKKFENWN
F3	MIGSIFYSKKFNLHRPSEYH
F4	MKDYRPLLGAIKVKGDNVFS
F5	MDIEVLRRLLERELSSEH
F6	Unable to be analyzed

Table 1 (continued)

Sample	N-Terminal Amino Acid Sequence
F7	PFEIVFEGAKEFAQLID

Example 4

Preparation of Cassette DNAs

[0083] Ten micrograms of *Pyrococcus furiosus* genomic DNA prepared in Example 1 was completely digested with *Eco*RI (manufactured by Takara Shuzo Co., Ltd.), and 500 ng equivalent of the digest was mixed with 50 ng of *Eco*RI cassette (manufactured by Takara Shuzo Co., Ltd.), followed by ligation. The DNA recovered from the ligation reaction mixture for ligation by ethanol precipitation was dissolved in 20 µl of sterilized water, and this solution was used as *Eco*RI cassette DNA for the subsequent procedures.

[0084] Using similar procedures as those described above, cassette DNAs ligated with each of *Hind*III cassette, *Xba*I cassette, *Sal*I cassette, *Pst*I cassette and *Sau*3AI cassette (all manufactured by Takara Shuzo Co., Ltd.) were prepared. When ligated with the *Xba*I cassette, genomic DNA digested with two enzymes, i.e., *Xba*I and *Nhe*I, was used, and each of the DNAs obtained were named *Xba*I cassette DNA and *NheI*/*Xba*I cassette DNA, respectively. When ligated with the *Sal*I cassette, genomic DNA digested with the two enzymes *Sal*I and *Xho*I was used, and each of the DNAs obtained were named *Sal*I cassette DNA and *XhoI*/*Sal*I cassette DNA, respectively. When ligated with the *Sau*3AI cassette, genomic DNA digested with *Bgl*II was used, and the DNA obtained was named *Bgl*II/*Sau*3AI cassette DNA.

Example 5

(1) Selection of Cosmid Clones Carrying F1 Gene

[0085] On the basis of the N-terminal amino acid sequence of F1 obtained in Example 3, the primers F1-1 and F1-2, of which nucleotide sequences are shown in SEQ ID NOs: 13 and 14, respectively, in Sequence Listing, were synthesized. First PCR was carried out using 100 pmol each of F1-1 and the cassette primer C1 (manufactured by Takara Shuzo Co., Ltd.) with 1 µl of the *Eco*RI cassette DNA prepared in Example 4 as a template. Second PCR was carried out using 100 pmol each of F1-2 and the cassette primer C2 (manufactured by Takara Shuzo Co., Ltd.) with 1 µl of the resulting reaction mixture obtained as above as a template. For the two PCRs, Pfu DNA polymerase (α-type enzyme, manufactured by STRATAGENE) was used. The reaction mixture composition and reaction conditions are shown below: The reaction mixture comprises 20 mM Tris-HCl, pH 8.2, 10 mM KCl, 20 mM $MgCl_2$, 6 mM $(NH_4)_2SO_4$, 0.2 mM each of dATP, dCTP, dGTP, dTTP, 1% Triton X-100, 0.01% BSA and 2.5 units of Pfu DNA polymerase (final volume being 100 µl), and the reaction was carried out in 30 cycles for the first PCR and in 25 cycles for the second PCR, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 45°C (30 seconds) - 72°C (2 minutes). The PCR using Pfu DNA polymerase described in the Examples below was also carried out using the same reaction mixture composition. An amplified DNA fragment of about 550 bp was subcloned into plasmid vector pUC119 (manufactured by Takara Shuzo Co., Ltd.), and its nucleotide sequence was determined. Thereafter, on the basis of the sequence determined, the primers F1S1 and F.S2, of which nucleotide sequences are shown in SEQ ID NOs: 15 and 16, respectively, in Sequence Listing, were then synthesized. PCR was carried out using these F1S1 and F1S2 with the cosmid DNA mentioned in Example 1 as a template, whereby selecting cosmid clones carrying the F1 gene. This PCR was carried out using TaKaRa PCR amplification kit (manufactured by Takara Shuzo Co., Ltd.) in accordance with the instructions attached. As a result, there were found that cosmid clone Nos. 22, 46, 61, 133, 178, 180, 210 and 317 carry the F1 gene.

(2) Subcloning of F1 Gene

[0086] PCR was carried out using 20 pmol each of F1S1 and the cassette primer C2, or each of F1S2 and the cassette primer C2, with 1 µl of the *Hind*III cassette DNA prepared in Example 4 as a template. The PCR was carried out with the same reaction mixture composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme in 50 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (3 minutes). As a result, a DNA fragment of 570 bp was amplified by F1S2 and the cassette primer C2, whereas no DNA was amplified by F1S1 and the cassette primer C2. This finding anticipated that the *Hind*III site is located immediately upstream of the initiation codon for the F1 gene and at a distance from the annealing position of F1S1 such that DNA

cannot be amplified by Pfu DNA polymerase. With this in mind, Cosmid Clone No. 61, randomly selected from the cosmid clones carrying the F1 gene, was digested with *Hind*III, and DNA fragments of not smaller than 1.5 Kb were isolated, and each was subcloned into plasmid vector pTV118N (manufactured by Takara Shuzo Co., Ltd.). PCR was carried out using F1S1 and F1S2 as primers with each recombinant plasmid obtained as a template, to examine for the presence of the F1 gene. As a result, it was found that a *Hind*III fragment of about 2 kb carries the F1 gene. A plasmid in which the F1 gene in this DNA fragment ligated to downstream of the *lac* promoter of pTV118N vector was named pF1-4-10. As to the DNA inserts contained in this plasmid, a restriction endonuclease map for *Nco*I, *Eco*RI, *Bam*HI, *Pst*I, *Sac*I and *Nde*I was prepared. The results as shown in Figure 2 were obtained.

(3) Determination of Nucleotide Sequence of DNA Fragment Carrying F1 Gene

[0087] There was determined by the dideoxy method the nucleotide sequence of the DNA insert in the plasmid pF1-4-10 and each plasmid obtained by cutting out the *Nco*I-*Hind*III, *Eco*RI-*Eco*RI, *Bam*HI-*Pst*I, *Eco*RI-*Hind*III, *Hind*III-*Eco*RI and *Hind*III-*Bam*HI fragments from the plasmid, and subcloning each of the resulting fragments into plasmid vector pTV119N (manufactured by Takara Shuzo Co., Ltd.). A sequence of 2,009 bp in the nucleotide sequences of the DNA insert in pF1-4-10 determined totally on the basis of these results combined together is as shown in SEQ ID NO: 17 in Sequence Listing. As a result of analyzing the nucleotide sequence, there was revealed an open reading frame comprising the N-terminal amino acid sequence of F1. The above sequence is shown in SEQ ID NO: 18 in Sequence Listing, and the amino acid sequence of the F1 translation product as deduced from the above sequence is shown in SEQ ID NO: 19 in Sequence Listing, respectively. This amino acid sequence was searched for homology to the amino acid sequences of known proteins. As a result, it was found to be homologous to the *Haemophilus influenzae*-derived single-stranded DNA-specific exonuclease [*Science*, **269**, 496-512 (1995)]. The homology was 23.2% for the first half and 24.3% for the last half.

(4) Construction of Plasmid for F1 Expression

[0088] PCR was carried out using the primer F1Nc, of which nucleotide sequence is shown in SEQ ID NO: 20 in Sequence Listing, and the above primer F1S2 with the plasmid pF1-4-10 described in Example 5(2) as a template. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase. Using 1 ng of template DNA and 20 pmol each of the two primers, the reaction was carried out in 25 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (2 minutes). A fragment obtained by digesting an amplified DNA fragment of about 460 base pairs with *Nco*I and *Bgl*III (both manufactured by Takara Shuzo Co., Ltd.) and a DNA fragment obtained by digesting the above plasmid pF1-4-10 with *Bgl*II and *Hind*III were together inserted between the *Nco*I and *Hind*III sites of plasmid vector pTV118N (manufactured by Takara Shuzo Co., Ltd.). This plasmid was named pF1Nc-2. Of the DNA insert in the plasmid, in the PCR-amplified region, the nucleotide sequence was confirmed by the dideoxy method that there is no mutation caused by PCR.

(5) Preparation of Purified F1 Authentic Sample

[0089] *Escherichia coli* JM109/pF1Nc2, *Escherichia coli* JM109 transformed with the plasmid pF1Nc-2 obtained in Example 5(4), was cultured for 16 hours in 2 liters of LB medium containing 100 μg/ml ampicillin. After harvesting the cells, 33 ml of a heat-treated supernatant was obtained in the same manner as Example 2(1). Next, this solution was applied to RESOURCE Q column (manufactured by Pharmacia), previously equilibrated with buffer D (50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol, 10% glycerol), and the applied solution was chromatographed using FPLC system (manufactured by Pharmacia). The elution was carried out on a linear concentration gradient of 0 to 500 mM NaCl. F1 was eluted at 340 mM NaCl.

[0090] After 10 ml of the enzyme solution obtained by collecting the F1 fraction was concentrated using Centriflow CF50, the resulting concentrate was subjected to exchange with buffer D using PD-10 column (manufactured by Pharmacia), and 3.5 ml of the solution was applied to HiTrap Blue column (manufactured by Pharmacia), previously equilibrated with the same buffer. Using FPLC system, the column was washed with buffer D, and thereafter F1 was eluted with buffer D containing 2 M NaCl. Five milliliters of this fraction was concentrated using Centricon-10, and 120 μl of the concentrate was applied to Superdex 200 gel filtration column (manufactured by Pharmacia), previously equilibrated with 50 mM Tris-HCl, pH 8.0, containing 2 mM 2-mercaptoethanol and 75 mM NaCl. The elution was carried out with the same buffer, and as a result, F1 was eluted at a position corresponding to a molecular weight of about 49 kilodaltons. This molecular weight corresponds to the case where F1 is present as a monomer.

(6) Determination of Exonuclease Activity

[0091] The 5' → 3' and 3' → 5' exonuclease activities of the purified F1 authentic sample were examined in the following manner.

[0092] First, plasmid vector pUC119 (manufactured by Takara Shuzo Co., Ltd.) was digested with *Ssp*I (manufactured by Takara Shuzo Co., Ltd.) and subjected to agarose gel electrophoresis, and a DNA fragment of 386 bp was recovered from the gel and purified. This DNA fragment was labeled at the 5'-terminus using [γ-^{32}P]-ATP (manufactured by Amersham) and polynucleotide kinase (manufactured by Takara Shuzo Co., Ltd.), and the ^{32}P-labeled DNA fragment obtained was used as a substrate for detecting the 5' → 3' exonuclease activity. In addition, plasmid vector pUC119 was digested with *Sau*3AI (manufactured by Takara Shuzo Co., Ltd.), and a DNA fragment of 341 bp obtained was recovered and purified in the same manner as above. Furthermore, this DNA fragment was ^{32}P-labeled at the 3'-terminus by the fill-in reaction using [α-^{32}P]-dCTP (manufactured by Amersham) and Klenow fragment (manufactured by Takara Shuzo Co., Ltd.) to yield a substrate for detecting the 3' → 5' exonuclease activity. The above two kinds of labeled DNAs were purified by gel filtration through NICK column (manufactured by Pharmacia) and used for the reaction described below.

[0093] Ten microliters of a reaction mixture (20 mM Tris-HCl, pH 7.7, 15 mM $MgCl_2$, 2 mM 2-mercaptoethanol) containing 2 ng of each of these labeled DNA fragments and 12.5 µg of digest obtained by completely digesting λ-DNA (manufactured by Takara Shuzo Co., Ltd.) with *Hae*III (manufactured by Takara Shuzo Co., Ltd.), and the above purified F1 authentic sample was prepared and reacted at 85°C for 2.5, 5 or 7.5 minutes, and thereafter ethanol precipitation was carried out to precipitate the DNA. By determining the radioactivity in this supernatant using a liquid scintillation counter, the amount of substrate decomposed by exonuclease activity was determined. In the determination of the 5' → 3' exonuclease activity, 50 fmol of the purified F1 authentic sample was added, and in the determination of the 3' → 5' exonuclease activity, 125 pmol of the purified F1 authentic sample was added. These results are shown in Figures 3 and 4, respectively.

[0094] Figure 3 shows the results for the determination of 5' → 3' exonuclease activity, and Figure 4 shows the results for determination of the 3' → 5' exonuclease. In the figures, the abscissa indicates reaction time, and the ordinate indicates the ratio of radioactivity released in the supernatant to that contained in the entire reaction mixture. In addition in the figures, solid circles indicate the results obtained with the purified F1 authentic sample of the present invention, and open circles indicate a blank reaction without adding the purified F1 authentic sample. As shown in the figures, the purified F1 authentic sample of the present invention possesses both 5' → 3' and 3' → 5' exonuclease activities. Also, from the above results it was demonstrated that the 5' → 3' exonuclease activity is about 500 times as great as the 3' → 5' exonuclease activity.

Example 6

(1) Selection of Cosmid Clones Carrying F2 Gene

[0095] On the basis of the N-terminal amino acid sequence of F2 obtained in Example 3, the primers F2-2 and F2-3, of which nucleotide sequences are shown in SEQ ID NOs: 21 and 22, respectively, in Sequence Listing, were synthesized. First PCR was carried out using 100 pmol of the primer F2-2 and 20 pmol of the cassette primer C1 with 1 µl of the *Xba*I cassette DNA prepared in Example 4 as a template. Second PCR was carried out using 100 pmol of the primer F2-3 and 20 pmol of the cassette primer C2 with 1 µl of the resulting reaction mixture obtained as above as a template. For the two PCRs, Pfu polymerase C was used. The reaction mixture composition and reaction conditions are shown below: The reaction mixture comprises 10 mM Tris-HCl, pH 9.2, 75 mM KCl, 3.5 mM $MgCl_2$, 0.4 mM each of dATP, dCTP, dGTP and dTTP, 0.1% Triton X-100, 0.01% BSA and 2.0 units of Pfu polymerase C (final volume being 100 µl), and the reaction was carried out in 30 cycles for the first PCR and 25 cycles for the second PCR, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 45°C (30 seconds) - 72°C (2 minutes). An amplified DNA fragment of about 250 bp was subcloned into plasmid vector pUC119, and its DNA sequence was determined. On the basis of the sequence determined, the primers F2S3 and F2S4, of which nucleotide sequences are shown in SEQ ID NOs: 23 and 24, respectively, in Sequence Listing, were then synthesized. PCR was carried out using these primers with the cosmid DNA prepared in Example 1 as a template, whereby selecting cosmid clones carrying the F2 gene. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme and 20 pmol each of the primers in 25 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (2 minutes). As a result, there was found that Cosmid Clone No. 172 carries the F2 gene.

(2) Subcloning of F2 Gene

[0096] PCR was carried out using 20 pmol each of F2S3 and the cassette primer C2 or each of F2S4 and the cassette primer C2 as primers with 1 μl of each of the *NheI/XbaI* and *XhoI/SalI* cassette DNAs of Example 4 as a template. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme in 50 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (3 minutes). As a result, each of amplified DNA fragments of about 700 bp and of about 1,400 bp for the *NheI/XbaI* and *XhoI/SalI* cassette DNAs, respectively, was amplified by the primer pair of F2S3 and the cassette primer C2, whereas no DNA was amplified by the primer pair of F2S4 and the cassette primer C2. This finding anticipated that the *NheI* and *XhoI* sites are located at a distance from the annealing position of the F2S4 primer unamplifiable with Pfu DNA polymerase.

[0097] With this in mind, the various DNA fragments obtained by digesting No. 172 with *NheI* were cut out, and each was suboloned into plasmid vector pTV118N (manufactured by Takara Shuzo Co., Ltd.). PCR was carried out using F2S3 and F2S4 as primers with each recombinant plasmid obtained as a template, to examine whether or not the F2 gene is present. As a result, it was found that an *NheI* fragment of about 8 kb carries the F2 gene. A plasmid resulting from insertion of this *NheI* fragment into pTV118N was named plasmid pF2172Nh. In addition, a restriction endonuclease map was prepared for the DNA insert in this plasmid. The results as shown in Figure 5 were obtained.

[0098] On the basis of the restriction endonuclease map shown in Figure 5, the plasmid pF2172Nh was digested with *HindIII*, and a *HindIII* fragment of about 1.5 kb was cut out, and each was subcloned into plasmid vector pTV118N. The recombinant plasmid obtained was examined for the insert orientation of the F2 gene, and there was found that the F2 gene was inserted in the reverse orientation with respect to the *lac* promoters of all of the vectors. This plasmid was named pF2172H16. *Escherichia coli* JM109/pF2172H16, *Escherichia coli* JM109 transformed with this plasmid, was examined for F2 expression, and found not to be highly expressed. With this in mind, in order to ligate the F2 gene in the orthodox orientation for the vector, pF2172H16 was digested with *HindIII* and *EcoRI*, and the *HindIII-EcoRI* fragment cut out was ligated to plasmid vector pTV119Nd (those resulting from substitution of the *NcoI* site with *NdeI* in plasmid vector pTV119N manufactured by Takara Shuzo Co., Ltd.). The recombinant plasmid obtained was named pF2172HE11, and *Escherichia coli* JM109 transformed with this plasmid was named *Escherichia coli* JM109/pF2172HE11.

(3) Preparation of F2 Authentic Sample

[0099] *Escherichia coli* JM109/pF2172HE11 obtained in Example 6(2) was cultured for 16 hours in 2 liters of LB medium containing 1 mM IPTG and 100 μg/ml ampicillin. After harvesting, cells were suspended in 23.4 ml of sonication buffer, and 19.5 ml of a heat-treated supernatant was obtained in the same manner as Example 2(1). Next, this solution was applied to RESOURCE Q column, previously equilibrated with buffer D, and the applied solution was chromatographed using FPLC system. F2 flowed through RESOURCE Q column.

[0100] Twenty-two milliliters of the flow-through F2 fraction was applied to RESOURCE S column (manufactured by Pharmacia), previously equilibrated with buffer D. Using FPLC system, the elution was carried out on a linear concentration gradient of 0 to 500 mM NaCl, and an F2 fraction was eluted at 170 mM NaCl. This fraction was concentrated using Centricon-10, and 75 μl of the concentrate obtained was applied to Superdex 200 gel filtration column, previously equilibrated with 50 mM Tris-HCl buffer (pH 8.0) containing 2 mM 2-mercaptoethanol and 75 mM NaCl. The elution was carried out with the same buffer, and as a result, F2 was eluted at a position corresponding to a molecular weight of about 120 kilodaltons or about 45 kilodaltons. This molecular weight corresponds to the case where F2 has formed a hexamer or dimer.

(4) Determination of Nucleotide Sequence of DNA Fragment Carrying F2 Gene

[0101] The nucleotide sequence of the DNA insert in the above plasmid pF2172HE11 was determined by the dideoxy method. A sequence of 957 bp of the nucleotide sequence determined is shown in SEQ ID NO: 25 in Sequence Listing. As a result of analyzing the nucleotide sequence, there was found an open reading frame having the N-terminal amino acid sequence of F2. The nucleotide sequence of this open reading frame is shown in SEQ ID NO: 26 in Sequence Listing, and the amino acid sequence of the F2 translation product as deduced from the nucleotide sequence is shown in SEQ ID NO: 27 in Sequence Listing, respectively. This amino acid sequence was searched for homology to the amino acid sequences of known proteins, and as a result, the homologous proteins were not found.

Example 7

(1) Selection of Cosmid Clones Carrying F4 Gene

[0102] On the basis of the N-terminal amino acid sequence of F4 obtained in Example 3, the primers F4-1 and F4-2, of which nucleotide sequences are shown in SEQ ID NOs: 28 and 29, respectively, in Sequence Listing, were synthesized. First PCR was carried out using 100 pmol of the primer F4-1 and 20 pmol of the cassette primer C1 with 1 μl of the *Hind*III cassette DNA of Example 4 as a template. Second PCR was carried out using F4-2 and the cassette primer C2 with 1 μl of the reaction mixture as a template. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme in 30 cycles for the first PCR and 25 cycles for the second PCR, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 45°C (30 seconds) - 72°C (2 minutes). An amplified DNA fragment of about 1,100 bp by this reaction was subcloned into plasmid vector pUC119, and a part of its nucleotide sequence was determined by the dideoxy method using M4 and RV primers (manufactured by Takara Shuzo Co., Ltd.). On the basis of the sequence determined, the primers F4S1 and F4S2, of which nucleotide sequences are shown in SEQ ID NOs: 30 and 31, respectively, in Sequence Listing, were then synthesized. PCR was carried out using these F4S1 and F4S2 primers with the cosmid DNA prepared in Example 1 as a template, whereby selecting cosmid clones carrying the F4 gene. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme in 30 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (1 minute). As a result, it was found that Cosmid Clone Nos. 16, 26, 88, 112, 250, 269, 427 and 451 carry the F4 gene.

(2) Subcloning of F4 Gene

[0103] PCR was carried out using 20 pmol each of F4S2 and the cassette primer C2 with 1 μl of the *Xba*I cassette DNA of Example 4 as a template. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme in 50 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (3 minutes). As a result, a DNA fragment of about 700 bp was amplified with F4S2 and the cassette primer C2. Also, PCR was carried out under the same conditions using F4-2 and the cassette primer C2 with *Hind*III cassette DNA as a template. As a result, a DNA fragment of about 1,100 bp was amplified. These findings suggested that the F4 gene is present in an *Xba*I-*Hind*III fragment of about 1.6 kb. With this in mind, Cosmid No. 16 was digested with *Xba*I and *Hind*III, and a DNA fragment of about 1.6 kb was cut out, and each was subcloned into pTV118N vector. PCR was carried out using the F4S1 and F4S2 primers with each recombinant plasmid obtained as a template, in order to examine for the presence of the F4 gene. As a result, a plasmid harboring a 1.6 kb *Xba*I-*Hind*III fragment carrying the F4 gene was obtained, and this plasmid was named plasmid pF4-1-4. Also, this plasmid was digested with the restriction enzymes *Nco*I, *Eco*RI, *Bam*HI, *Pst*I, *Sac*I and *Nde*I. As a result, it was found that none of these sites were present in the above plasmid or DNA insert.

(3) Determination of Nucleotide Sequence of DNA Fragment Carrying F4 Gene

[0104] The nucleotide sequence of the DNA insert in the above plasmid pF4-1-4 was determined by the dideoxy method.

[0105] A sequence of 1,012 bp of the nucleotide sequence determined is shown in SEQ ID NO: 32 in Sequence Listing. As a result of analyzing the nucleotide sequence, there was found an open reading frame having the N-terminal amino acid sequence of F4. The nucleotide sequence of this open reading frame is shown in SEQ ID NO: 33 in Sequence Listing, and the amino acid sequence of the F4 translation product as deduced from the nucleotide sequence is shown in SEQ ID NO: 34 in Sequence Listing, respectively. This amino acid sequence was searched for homology to the amino acid sequences of known proteins, and as a result, the homologous proteins were not found.

(4) Construction of Plasmid for F4 Expression

[0106] PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) with Pfu DNA polymerase using the primer F4NNd, of which nucleotide sequence is shown in SEQ ID NO: 35 in Sequence Listing, and the primer F4CEc, of which nucleotide sequence is shown in SEQ ID NO: 36 in Sequence Listing, with the plasmid pF4-1-4 described in Example 7(3) as a template. The reaction conditions are shown below. Using 1 ng of template DNA and 20 pmol each of the two primers, the reaction was carried out in 25 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (2 minutes). An amplified DNA fragment of about 450 bp was digested with *Nde*I and *Eco*RI (both manufactured by Takara Shuzo Co., Ltd.), and the DNA fragment obtained was inserted between the *Nde*I and *Eco*RI sites of plasmid vector pTV119Nd mentioned above to prepare the

plasmid pF4Nd-6. Furthermore, the nucleotide sequence of the DNA insert in the plasmid was determined by the dideoxy method. It was confirmed that there is no mutation caused by PCR.

(5) Preparation of Purified F4 Authentic Sample

[0107] *Escherichia coli* JM109/p4Nd-6, *Escherichia coli* JM109 transformed with the plasmid pF4Nd-6 obtained in Example 7(4), was cultured for 16 hours in 2 liters of LB medium containing 100 μg/ml ampicillin. After harvesting, cells were suspended in 33.4 ml of sonication buffer, and 28 ml of a heat-treated supernatant was obtained in the same manner as Example 2(1). Next, this solution was applied to RESOURCE Q column, previously equilibrated with buffer D, and the applied solution was chromatographed using FPLC system. The elution was carried out on a linear concentration gradient of 0 to 500 mM NaCl. F4 was eluted at a concentration of 325 mM NaCl.

[0108] Three milliliters of the solution obtained by collecting the F4 fraction was subjected to exchange with buffer D containing 150 mM NaCl using PD-10 column, and 6.9 ml of the solution was applied to HiTrap Heparin column, previously equilibrated with the same buffer. F4 was not adsorbed to HiTrap Heparin column, and $(NH_4)_2SO_4$ was added to 7.2 ml of the F4 fraction flowed through the column so as to have a final concentration of 1 M. This solution was applied to HiTrap Phenyl column (manufactured by Pharmacia), previously equilibrated with buffer D containing 1 M $(NH_4)_2SO_4$. Using FPLC system, the column was washed with each of 1 M and 0.5 M $(NH_4)_2SO_4$, and thereafter F4 was eluted with buffer D. Five milliliters of this fraction was concentrated using Centricon-10, and 76 μl of the concentrate obtained was applied to Superdex 200 gel filtration column, previously equilibrated with 50 mM Tris-HCl buffer, pH 8.0, containing 2 mM 2-mercaptoethanol and 75 mM NaCl. As a result of the elution with the same buffer, F4 was eluted at a position corresponding to a molecular weight of about 39 kilodaltons. This molecular weight corresponds to the case where F4 has formed a dimer or trimer.

Example 8

(1) Selection of Cosmid Clones Carrying F7 Gene

[0109] On the basis of the N-terminal amino acid sequence of F7 obtained in Example 3, the primers F7-1 and F7-2, of which nucleotide sequences are shown in SEQ ID NOs: 37 and 38, respectively, in Sequence Listing, were synthesized. First PCR was carried out using 100 pmol of F7-1 and 20 pmol of the cassette primer C1 with 1 μl of the *Hind*III cassette DNA prepared in Example 4 as a template. Second PCR was carried out using 100 pmol of the primer F7-2 and 20 pmol of the cassette primer C2 with 1 μl of the reaction mixture obtained as above as a template. The PCR was carried out using the same reaction mixture composition and reaction conditions as those used in Example 6(1). An amplified DNA fragment of about 830 bp was subcloned into plasmid vector pUC119, and its nucleotide sequence was determined. On the basis of the sequence determined, the primers F7S1 and F7S2, of which nucleotide sequences are shown in SEQ ID NOs: 39 and 40, respectively, in Sequence Listing, were then synthesized. PCR was carried out using these primers with the cosmid DNA described in Example 1 as a template, whereby selecting cosmid clones carrying the F7 gene. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme in 30 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (3 minutes). As a result, there was found that Cosmid Clone Nos. 15, 96, 114, 167, 277, 348, 386, 400, 419, 456, 457 and 484 carry the F7 gene.

(2) Subcloning of F7 Gene

[0110] PCR was carried out using 20 pmol each of F7S2 and the cassette primer C2 with 1 μl of the *Hind*III cassette DNA prepared in Example 4 as a template. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme in 50 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (3 minutes). As a result, a fragment of about 900 bp was amplified. From this result, together with the result of amplification using F7-2 of Example 8(1) and the cassette primer C2, the presence of the F7 gene in a *Hind*III fragment of about 1.0 kb was anticipated. With this in mind, No. 15, randomly selected from the cosmids carrying that gene, was digested with *Hind*III, and a DNA fragment of around 1.0 kb was cut out, and each was subcloned into plasmid vector pTV118N. PCR was carried out using the F7S1 and F7S2 primers with each recombinant plasmid obtained as a template, to examine for the presence of the F7 gene, and as a result, it was found that a *Hind*III fragment of 1 kb carries the F7 gene. A plasmid in which the F7 gene in this DNA fragment was ligated to downstream of the *lac* promoter of pTV118N vector was named pF7-HH-18, and a plasmid in which the F7 gene was ligated in the opposite orientation was named pF7-1-8. Also, a restriction endonuclease map was prepared for the DNA insert contained in this plasmid, and the map as shown in Figure 6 was obtained.

(3) Determination of Nucleotide Sequence of DNA Fragment Carrying F7 Gene

[0111] There was determined by the dideoxy method the nucleotide sequence of each insert in the above two kinds of plasmids, each insert in the plasmids being prepared by cutting out the *Bam*HI-*Hind*III, *Nde*I-*Hind*III, *Hind*III-*Nde*I and *Hind*III-*Bam*HI fragments from the above two kinds of plasmids, and subcloning the fragments into plasmid vector pTV119Nd. A sequence of 989 bp of the nucleotide sequence of the DNA insert of the above plasmid, determined on the basis of these overall results, is shown in SEQ ID NO: 41 in Sequence Listing. As a result of analyzing the nucleotide sequence, there was found an open reading frame containing the N-terminal amino acid sequence of F7. The nucleotide sequence of this open reading frame is shown in SEQ ID NO: 2 in Sequence Listing, and the amino acid sequence of the F7 translation product as deduced from the nucleotide sequence is shown in SEQ ID NO: 1 in Sequence Listing. This amino acid sequence was searched for homology to the amino acid sequences of known proteins, and as a result, it was found that the amino acid sequence was homologous to the proliferating cell nuclear antigen (PCNA) involved in the DNA replication in eukaryotes [*EMBO J.*, **11**, 5111-5120 (1995); *Nucleic Acids Research*, **18**, 261-265 (1990); *Proc. Natl. Acad. Sci. USA*, **84**, 1575-1579 (1987)]. The homology to the proteins described in the individual references were 24, 28 and 24%, respectively.

(4) Preparation of Purified F7 Authentic Sample

[0112] *Escherichia coli* JM109/pF7-HH-18, *Escherichia coli* JM109 transformed with the plasmid pF7-HH-18 obtained in Example 8(2), was cultured for 16 hours in 2 liters of LB medium containing 100 μg/ml ampicillin. After harvesting, cells were suspended in 45 ml of sonication buffer, and 41.9 ml of a heat-treated supernatant was obtained in the same manner as Example 2(1). Next, this solution was thrice subjected to 2-hour dialysis against 2 liters of buffer A as a dialysate. After dialysis, 36 ml of the enzyme solution was applied to RESOURCE Q column, previously equilibrated with buffer A, and the applied solution was chromatographed using FPLC system. The elution was carried out on a linear concentration gradient of 0 to 500 mM NaCl. As a result, F7 was eluted at 340 mM NaCl.

[0113] Ten milliliters of the solution obtained by collecting the F7 fraction was concentrated using Centriflow CF-50, and thereafter subjected to exchange with buffer A containing 1 M $(NH_4)_2SO_4$ using PD-10 column, and 3.5 ml of the solution obtained was applied to HiTrap Phenyl column, previously equilibrated with the same buffer. Using FPLC system, the column was sequentially washed with 1 M and 0.5 M $(NH_4)_2SO_4$, and thereafter F7 was eluted with buffer A. Four milliliters of this fraction was concentrated using Centricon-10, and 80 μl of this concentrate was applied to Superdex 200 gel filtration column, previously equilibrated with 50 mM potassium phosphate buffer (pH 6.5) containing 2 mM 2-mercaptoethanol and 75 mM NaCl. As a result of elution with the same buffer, F7 was eluted at a position corresponding to a molecular weight of about 99 kilodaltons. This molecular weight corresponds to the case where F7 has formed a trimer.

(5) Effects of F7 on Primer Extension Reactions

[0114] In order to examine for the effects of F7 on the primer extension reactions to various polymerases, the activities of Pfu polymerase C, Pfu DNA polymerase (α-type DNA polymerase, manufactured by STRATAGENE) and *Pyrodictium occultum*-derived Poc DNA polymerases I and II [Poc DNA polymerases I and II, *J. Bacteriol.*, **177**, 2164-2177 (1995)] were compared with regard to the presence or absence of the addition of F7.

[0115] Determination of DNA polymerase activities were carried out with reference to the Pfu polymerase C activity determination described in Example 2(1). The substrate used was the constructs (M13-HT primer) as prepared by annealing the HT primer, a synthetic oligonucleotide having 45 bases, to M13 phage single-stranded DNA (M13mp18 ssDNA, manufactured by Takara Shuzo Co., Ltd.). The nucleotide sequence of the HT primer is shown in SEQ ID NO: 42 in Sequence Listing.

[0116] Concretely, a reaction mixture [20 mM Tris-HCl, pH 7.7, 15 mM $MgCl_2$, 2 mM 2-mercaptoethanol, 0.01 μg/μl M13-HT primer, 40 μM each of dATP, dCTP, dGTP and dTTP, 60 nM [^{3}H]-dTTP (manufactured by Amersham)] making up a final volume of 50 μl and containing each DNA polymerase listed in Table 2 and F7 was prepared and reacted at 75°C for 5 minutes. After the reaction mixture was cooled with ice to stop the reaction, a 40 μl portion was spotted onto DE paper (manufactured by Whatman) and washed 5 times with 5% Na_2HPO_4, and thereafter the remaining radioactivity on the DE paper was determined using a liquid scintillation counter.

[0117] As shown in Table 2, for all the DNA polymerases used, an increase in DNA polymerase activity due to the addition of F7 was observed.

Table 2

DNA Polymerase		F7	Enzyme Activity (cpm)
Blank 1		-	61
Blank 2		10pmol	35
Pfu Polymerase C	(25fmol)	-	888
Pfu Polymerase C	(25fmol)	5pmol	2897
Pfu Polymerase C	(25fmol)	10pmol	3175
Pfu DNA Polymerase	(120fmol)	-	907
Pfu DNA Polymerase	(120fmol)	0.48pmol	1363
Pfu DNA Polymerase	(120fmol)	4.8pmol	1637
Poc DNA Polymerase I	(74pmol)	-	62
Poc DNA Polymerase I	(74pmol)	10pmol	69
Poc DNA Polymerase II	(6.0pmol)	-	433
Poc DNA Polymerase II	(6.0pmol)	10pmol	1443
Note: In the table, the amount of Pfu polymerase C is the amount of a protein comprising one molecule each of the two DNA polymerase-constituting proteins, and the amount of F7 is the amount as a trimer protein.			

[0118] Primer extension activity was further studied in detail. The M13-HT primer, previously labeled at the 5'-terminus of the primer using [γ-^{32}P]-ATP (manufactured by Amersham) and T4 polynucleotide kinase (manufactured by Takara Shuzo Co., Ltd.), was used as a substrate.

[0119] A 1 µl sample solution containing each of the following samples was prepared: 1) 18 fmol of Pfu polymerase C, 2) 18 fmol of Pfu polymerase C + 2 pmol of F7, 3) 0.24 pmol of Pfu DNA polymerase, 4) 0.24 pmol of Pfu DNA polymerase + 0.78 pmol of F7. To each sample solution, 9 µl of a reaction mixture [20 mM Tris-HCl (pH 9.0), 15 mM $MgCl_2$, 2 mM 2-mercaptoethanol, 40 µM each of dATP, dGTP, dCTP and dTTP] containing 0.01 µg/µl ^{32}P-labeled M13-HT primer was added, and a reaction was carried out at 75°C for 2.5 minutes or 5 minutes. After termination of the reaction, the reaction mixture was cooled with ice to stop the reaction, and 1 µl of 200 mM EDTA and 5.5 µl of a reaction stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) were added, and thermal denaturation treatment was carried out at 95°C for 5 minutes. After 1.6 µl of this reaction mixture was electrophoresed using 6% polyacrylamide gel containing 8 M urea, an autoradiogram was prepared. The autoradiogram obtained is shown in Figure 7.

[0120] In the figure, Pfu-C and pfu show the results obtained with Pfu polymerase C and Pfu DNA polymerase, respectively, and 2.5 and 5 show the respective reaction time (minutes). In addition, the symbols - and + in the figure show the results obtained with the reaction mixture in the absence and presence of F7, respectively. Further, the lanes on both ends of the figure show the results of electrophoresis of λ-*Eco*T14I digest (manufactured by Takara Shuzo Co., Ltd.), previously labeled at the 5'-terminus using [γ-^{32}P]-ATP (manufactured by Amersham) and T4 polynucleotide kinase (manufactured by Takara Shuzo Co., Ltd.), and were used to deduce the lengths of the extension products.

[0121] As shown in Figure 7, when F7 is not added, in Pfu polymerase C, DNAs of about 300 to 600 bases are the major extension products obtained, whereas when F7 is added, extension products of low chain length decreases and the ratio of extension products exceeding 1,000 bases increases. Also in Pfu DNA polymerase, the chain length of extension products was markedly extended by the addition of F7. It was thus elucidated that F7 increases the primer extension rates of both Pfu polymerase C and Pfu DNA polymerase.

[0122] Next, in order to analyze primer extension reaction products of higher molecular weights, the primer extension reaction products of Pfu polymerase C and Pfu DNA polymerase with the ^{32}P-labeled M13-HT primer as a substrate were analyzed by alkaline agarose gel electrophoresis. To 1 µl of a solution of each of samples 1) to 4) above, 9 µl of a reaction mixture (20 mM Tris-HCl, pH 9.0, 15 mM $MgCl_2$, 2 mM 2-mercaptoethanol, 40 µM each of dATP, dGTP, dCTP and dTTP, 84 nM [α-^{32}P]-dCTP) was added so as to have a final concentration of 0.01 µg/µl M13-HT primer, and a reaction was carried out at 75°C for 2.5 minutes. After termination of the reaction, to the ice cooled reaction mixture, 1.11 µl of 200 mM EDTA, 1.23 µl of 500 mM NaOH and 2.47 µl of 6-fold concentrated loading buffer (0.125% bromophe-

nol blue, 0.125% xylene cyanol, 9% glycerol) were sequentially added. After 6 µl of this mixture was electrophoresed using 0.5% alkaline agarose gel, an autoradiogram was prepared. The autoradiogram obtained is shown in Figure 8.

[0123] In the figure, Pfu-C and pfu show the results obtained with Pfu polymerase C and Pfu DNA polymerase, respectively, and the symbols - and + in the figure show the results obtained without or with addition of F7, respectively. Further, in the figure, Lane M is for the λ-*Eco*T14I digest, previously labeled at one end in the same manner as above. As shown in Figure 8, in the case of Pfu polymerase C, a weak extension product signal was observed near 2.5 kb in the absence of F7, whereas a 7.3 kb signal completely encircling M13 ssDNA was observed in the presence of F7. In addition, in the case of Pfu DNA polymerase, a signal was observed near 2.7 kb in the presence of F7, whereas no signal was observed in the absence of F7. These findings demonstrate that F7 enhances the extension reactions of the two DNA polymerases.

Example 9

(1) Selection of Cosmid Clones Carrying Gene Encoding Homologs of RFC Small Subunit

[0124] Regarding the amino acid sequence of the RFC small subunit of *Methanococcus jannaschii* [Science, **273**, 1058-1073 (1996)], homology to the amino acid sequences of RFC (RF-C) small subunits derived from other organisms was examined. On the basis of the amino acid sequences of regions highly conserved thereamong, the primers RF-F1, RF-F3, RF-F4, RF-R1, RF-R2, RF-R3 and RF-R4 for searching the gene encoding the RFC small subunit were synthesized. The nucleotide sequences of these primers are shown in SEQ ID NOs: 43 to 49, respectively, in Sequence Listing. PCR was carried out using various combinations of these primers with *Pyrococcus furiosus* genomic DNA as a template, whereby searching for the gene encoding the RFC small subunit. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase, and using 0.25 µg of template DNA and 100 pmol of each primer. When first PCR was carried out using RF-F1 and RF-R4, second PCR was carried out using RF-F4 and RF-R4, or RF-F1 and RF-R1, with 1 µl of the reaction mixture as a template. When first PCR was carried out using RF-F1 and RF-R3, second PCR was carried out using RF-F3 and RF-R2 with 1 µl of the reaction mixture as a template. Amplified DNA fragments of about 240 bp, about 140 bp and about 140 bp, respectively, were obtained. Each of these DNA fragments was subcloned into plasmid vector pUC119, and its nucleotide sequence was determined. On the basis of the sequences determined, the primers RF-S1, RF-S2, RF-S3, RF-S4 and RF-S5, of which nucleotide sequences are shown in SEQ ID NOs: 50 to 54, respectively, in Sequence Listing, were then synthesized. PCR was carried out using these RF-S1 and RF-S3 primers with the cosmid DNA prepared in Example 1 as a template, whereby selecting cosmid clones assumed to carry the gene encoding homologs of the RFC small subunit. The PCR was carried out using the TaKaRa PCR amplification kit in 25 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (2 minutes). As a result, there was found that Cosmid Clone Nos. 254, 310, 313, 377 and 458 carry the desired gene (PFU-RFC gene).

(2) Subcloning of PFU-RFC Gene

[0125] PCR was carried out using 100 pmol of RF-S1 and 20 pmol of the cassette primer C2, or 100 pmol of RF-S2 and 20 pmol of the cassette primer C2, with 1 µg each of the *Xba*I and *Eco*RI cassette DNAs prepared in Example 4 as a template. The PCR was carried out in a reaction mixture of the same composition as that used in Example 6(1) using the Pfu polymerase C enzyme in 50 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (3 minutes). As a result, a DNA fragment of about 2 kb was amplified by RF-S1 and the cassette primer C2 when the *Xba*I cassette was used as a template, and a DNA fragment of about 1.5 kb was amplified by RF-S2 and the cassette primer C2 when the *Eco*RI cassette was used as a template. Each of these DNA fragments was subcloned into plasmid vector pUC119, and the recombinant plasmids obtained were named pRFSXS1-26 and pRFSES2-8. Restriction endonuclease maps of these plasmids were prepared, and as a result, it was anticipated that neither *Nde*I nor *Bam*HI site is present in the PFU-RFC gene.

[0126] The cosmids of the five clones mentioned in (1) above were each digested with *Nde*I and *Bam*HI, and the electrophoretic patterns were examined. As a result, a common band was observed near 5 kb. Anticipating the presence of the PFU-RFC gene in this DNA fragment, an *Nde*I-*Bam*HI fragment of about 5 kb from Clone No. 254 was cut out, and each was subcloned into pTV119Nd vector mentioned above. A transformant formed with the recombinant plasmid obtained was examined for the presence PFU-RFC gene by PCR using the RF-S1 and RF-S3 primers. As a result, there was found that this *Nde*I-*Bam*HI fragment carry the PFU-RFC gene. Therefore, the plasmid resulting from insertion of this *Nde*I-*Bam*HI fragment into pTV119Nd vector was named plasmid pRFS254NdB. In addition, a restriction endonuclease map of this plasmid was prepared, and the map as shown in Figure 9 was obtained.

[0127] On the basis of the restriction endonuclease map shown in Figure 9, various fragments were cut out from pRFS254NdB by the method described below, and each was subcloned into pTV118N vector (manufactured by Takara

Shuzo Co., Ltd.). First, a DNA fragment of about 500 bp obtained by digesting pRFS254NdB with *Xba*I and *Sac*I, a DNA fragment of about 2 kb obtained by digesting with *Xba*I and *Nco*I, and a DNA fragment of about 1.1 kb obtained by digesting with *Nco*I and *Bam*HI was prepared, respectively, and each was mixed with pTV118N, previously linearized with *Sac*I and *Bam*HI, for ligation, whereby constructing a recombinant plasmid. This plasmid was named pRFS254SXNB.

(3) Determination of Nucleotide Sequence of DNA Fragment Carrying PFU-RFC Gene

[0128] The nucleotide sequence of the DNA insert in the plasmid pRFS254NdB obtained in Example 9(2) was determined by the dideoxy method. A sequence of 3,620 base pairs of the nucleotide sequence determined is shown in SEQ ID NO: 55 in Sequence Listing. The amino acid sequence of the protein encoded by this nucleotide sequence was deduced. As a result of comparing this amino acid sequence with those of known RFC small subunits, there was anticipated the presence of one intein in the amino acid sequence of PFU-RFC. This intein is encoded by Nos. 721 to 2295 of SEQ ID NO: 55 in Sequence Listing.

(4) Construction of Intein-Eliminated PFU-RFC Expression Plasmid

[0129] On the basis of the nucleotide sequence determined in Example 9(3), and the amino acid sequence of a known RFC small subunit and the nucleotide sequence of the gene encoding the subunit, the primers RF-CBΔI and RF-CAΔI, of which nucleotide sequences are shown in SEQ ID NOs: 56 and 57 in Sequence Listing, were synthesized. Inverse PCR was carried out using these two primers, each of which 5'-terminus was previously phosphorylated, with the above plasmid pRFS254SXNB as a template. For inverse PCR, TaKaRa Ex Taq was used to prepare 100 μl of a reaction mixture in accordance with the instructions for the enzyme. To this reaction mixture added with 15 ng of the plasmid pRFS254SXNB and 20 pmol each of the primers, the reaction was carried out in 30 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (3 minutes). An amplified DNA fragment obtained by the inverse PCR was blunt-ended using DNA blunting kit (manufactured by Takara Shuzo Co., Ltd.), and thereafter subjected to self-ligation, whereby constructing a plasmid, which was named the plasmid pRFS254ISΔI.

[0130] Furthermore, an *Xba*I-*Nco*I fragment of about 400 bp isolated after digestion of the plasmid with *Xba*I and *Nco*I was mixed with and an *Xba*I-*Sac*I fragment of about 500 bp and an *Nco*I-*Bam*HI fragment of about 1.1 kb, each isolated from the plasmid pRFS254NdB obtained in Example 9(2), and the mixed fragments were subcloned between the *Bam*HI and *Sac*I sites of plasmid vector pTV118N. The recombinant plasmid obtained as described above was named pRFS254SNc. *Escherichia coli* JM109 transformed with the plasmid was named *Escherichia coli* JM109/pRFS254SNC. It was found that the transformant expresses PFU-RFC at high level.

(5) Determination of Nucleotide Sequence of Gene Encoding PFU-RFC Without Carrying Intein

[0131] An *Xba*I-*Nco*I fragment of about 400 bp derived from the plasmid pRFS254SXNB obtained in Example 9(4) was subcloned into plasmid vector pTV118N, and the nucleotide sequence of the DNA insert was determined, whereby the nucleotide sequence encoding the boundary portion of the intein eliminated was confirmed. From this result and the results of Example 9(3), the nucleotide sequence of the gene encoding PFU-RFC without carrying intein was determined. The nucleotide sequence of the open reading frame encoding PFU-RFC without carrying intein obtained as described above and the amino acid sequence of PFU-RFC deduced from the nucleotide sequence are shown in SEQ ID NOs: 4 and 3, respectively, in Sequence Listing.

(6) Preparation of Purified PFU-RFC Authentic Sample

[0132] *Escherichia coli* JM109/pRFS254Nc obtained in Example 9(4) was cultured for 16 hours in 2 liters of LB medium containing 100 μg/ml ampicillin. After harvesting, cells were suspended in 44.1 ml of sonication buffer, and 35.2 ml of a heat-treated supernatant was obtained in the same manner as Example 2(1). Next, this solution was applied to RESOURCE Q column, previously equilibrated with buffer D, and the applied solution was chromatographed using FPLC system. PFU-RFC was flowed through RESOURCE Q column.

[0133] Thirty-five milliliters of the flow-through PFU-RFC fraction was applied to RESOURSE S column (manufactured by Pharmacia), previously equilibrated with buffer D. Using FPLC system, the elution was carried out on a linear concentration gradient of 0 to 500 mM NaCl to yield a PFU-RFC fraction eluted at 170 mM NaCl. 2.9 ml Of this fraction was concentrated using Centricon-10, and 105 μl of the concentrate obtained was applied to Superdex 200 gel filtration column, previously equilibrated with 50 mM Tris-HCl buffer, pH 8.0, containing 2 mM 2-mercaptoethanol and 75 mM NaCl. The elution was carried out with the same buffer, and as a result, PFU-RFC was eluted at a position corresponding to a molecular weight of about 150 kilodaltons. This molecular weight corresponds to the case where PFU-RFC has

formed a tetramer.

(7) Effects of PFU-RFC on Primer Extension Reaction

[0134] The effects of PFU-RFC and F7 on the primer extension reaction by Pfu polymerase C were examined in the same manner as Example 8(5). The results are shown in Table 3. As shown in Table 3, PFU-RFC slightly enhanced the activity of Pfu polymerase C. Furthermore, in the case where PFU-RFC was added simultaneously with F7, the enhanced activity more than doubled than the case where F7 was added alone.

Table 3

Pfu Polymerase C	F7	PFU-RFC	Enzyme Activity (cpm)
-	-	-	100
90 fmol	-	-	366
90 fmol	9.6 pmol	-	2743
90 fmol	-	356 fmol	463
90 fmol	9.6 pmol	356 fmol	8740

Note: In the table, the amount of Pfu polymerase C is the amount as a protein comprising one molecule each of the two DNA polymerase-constituting proteins, and the amounts of F7 and PFU-RFC are the amounts as a trimer and tetramer proteins, respectively.

Example 10

(1) Preparation of Anti-Pfu DNA Polymerase Antibody

[0135] Twelve milliliters (30,000 units) of cloned Pfu DNA polymerase (manufactured by STRATAGENE) was concentrated by ultrafiltration using Centricon-10, and thereafter 0.1 ml of the concentrate obtained was applied to Superdex 200 gel filtration column (manufactured by Pharmacia), previously equilibrated with 50 mM Tris-HCl (pH 8.0) containing 2 mM 2-mercaptoethanol and 75 mM NaCl. The elution was carried out with the same buffer, and a Pfu DNA polymerase fraction eluted at a position corresponding to a molecular weight of about 76 kilodaltons was recovered. After 0.8 ml of this fraction was concentrated using Centricon-10, this concentrate was used as an antigen to prepare an anti-Pfu DNA polymerase polyclonal antibody. The above concentrate was diluted with physiological saline so as to have a Pfu DNA polymerase concentration of 2 mg/ml, and the diluted solution was emulsified with an equal volume of Freund's complete adjuvant. This emulsion was subcutaneously injected to rabbits at 250 μl per injection four times at 3-week intervals. Ten days after final immunization, whole blood was extracted. After allowing to stand at room temperature for 60 minutes, the extracted blood was centrifuged to yield 60 ml of an antiserum containing the anti-Pfu DNA polymerase polyclonal antibody. To 26 ml of this antiserum, 26 ml of a saturated solution of ammonium sulfate was added, and the mixture was gently stirred at 4°C for 1 hour and 45 minutes, and subsequently centrifuged. The precipitate was suspended in 5 ml of 20 mM sodium phosphate buffer (pH 7.0) and desalted using PD-10 column (manufactured by Pharmacia), previously equilibrated with the same buffer. Ten milliliters of this solution was applied to Protein A column (manufactured by Pharmacia), previously equilibrated with 20 mM sodium phosphate buffer (pH 7.0). After the column was washed with the same buffer, the elution was carried out with 0.1 M sodium citrate buffer (pH 3.0). The eluted fraction containing the anti-Pfu DNA polymerase polyclonal antibody was neutralized with 1 M Tris-HCl, pH 9.0, and thereafter the mixture was concentrated using Centriflow CF-50 and subjected to exchange with coupling buffer (0.5 M NaCl, 0.2 M $NaHCO_3$, pH 8.3) using PD-10 column to prepare a solution containing the anti-Pfu DNA polymerase antibody.

(2) Preparation of Anti-Pfu DNA Polymerase Antibody Column

[0136] HiTrap NHS-activated column (manufactured by Pharmacia) was washed with 6 ml of 1 mM HCl, and thereafter 0.9 ml of the above anti-Pfu DNA polymerase polyclonal antibody solution (containing 4.5 mg equivalent of the anti-Pfu DNA polymerase antibody) was applied. Subsequently, an anti-Pfu DNA polymerase antibody column was prepared in the same manner as Example 2(3).

(3) Confirmation of Formation of Complex of Pfu DNA Polymerase and F7 Using Anti-Pfu DNA Polymerase Antibody Column

[0137] *Pyrococcus furiosus* DSM3638 was cultured in the same manner as the method described in Example 1 to yield cells in 9 liters of a culture medium. These cells were suspended in 33 ml of buffer C (50 mM Tris-HCl, pH 8.0, 0.1 mM ATP) containing 2 mM PMSF, and the resulting suspension was treated with an ultrasonic disrupter. The disrupted solution obtained was centrifuged at 12,000 rpm for 10 minutes, and 44 ml of the supernatant obtained was applied to the anti-Pfu DNA polymerase antibody column, previously equilibrated with buffer C. The column was washed with buffer C containing 0.1 M NaCl, and thereafter the Pfu DNA polymerase complex was eluted with elution buffer (50 mM Tris-HCl, pH 8.0, 8 M urea). This eluate was subjected to SDS-PAGE (12.5% polyacrylamide gel; 25 mM Tris-HCl, 192 mM glycine, 0.1% SDS, pH 8.4 used as electrophoresis buffer). The gel after electrophoresis was stained with Coomassie brilliant blue R-250 by a conventional method. As a result, as shown in Figure 10, besides the band of Pfu DNA polymerase, a band was detected at a position corresponding to the above F7.

[0138] With this in mind, a concentrate of this eluate was subjected to SDS-PAGE in the same manner as above, and the gel obtained was subjected to Western blotting using the anti-Pfu DNA polymerase antibody in the same manner as Example 3(2). From the result of SDS-PAGE shown in Figure 10 and the results of the above Western blotting, there was elucidated that the band at a position corresponding to F7 is a protein unreactive with the anti-Pfu DNA polymerase antibody.

[0139] Furthermore, the N-terminal amino acid sequence of the protein of this band was analyzed in the same manner as Example 3(2), and as a result, it was found that this protein is F7.

(4) Confirmation of Formation of Complex of Pfu DNA Polymerase and F7 Using Gel Filtration Chromatography

[0140] 1.2 ml Of the F7 authentic sample obtained in Example 8(4) was subjected to buffer-exchange with 50 mM Tris-HCl (pH 8.0) containing 2 mM 2-mercaptoethanol and 75 mM NaCl using PD-10 column, and thereafter the resulting solution was concentrated to a volume of 50 µl using Centricon-10.

[0141] Ten microliters each of the 0.1 mM Pfu DNA polymerase solution described in Example 10(1), the above 0.1 mM (calculated as a trimer) F7 solution, and a mixture of 0.1 mM Pfu DNA polymerase and 0.1 mM F7, was heated from 60° to 90°C over a period of 30 minute. Each heat-treated solution was applied to Superdex 200 PC3.2/30 gel filtration column (manufactured by Pharmacia), previously equilibrated with 50 mM Tris-HCl buffer, pH 8.0, containing 2 mM 2-mercaptoethanol and 75 mM NaCl, and the elution was carried out with the same buffer. Pfu DNA polymerase and F7 were eluted at positions corresponding to molecular weights of about 76 kilodaltons and about 128 kilodaltons, respectively. In the case of the mixture of Pfu DNA polymerase and F7, a main peak corresponding to about 320 kilodaltons and a minor peak corresponding to about 128 kilodaltons were eluted. The fractions with these two peaks were each subjected to SDS-PAGE (12.5% polyacrylamide gel; 25 mM Tris-HCl, 192 mM glycine, 0.1% SDS, pH 8.4 used as electrophoresis buffer). The fraction corresponding to about 320 kilodaltons contained Pfu DNA polymerase and F7, whereas the fraction corresponding to about 128 kilodaltons contained F7 only. From the above, there was found that a complex of Pfu DNA polymerase and F7 is formed.

(5) Extension Activity of Pfu DNA Polymerase-F7 Complex

[0142] In the gel filtration described in Example 10(4), 20 µl each of the eluates obtained by gel filtration of Pfu DNA polymerase alone corresponding to about 76 kilodaltons, and of the mixture of Pfu DNA polymerase and F7 corresponding to 320 kilodaltons, were each collected, and the primer extension activity of each eluate or mixture was determined by the activity determination method described in Example 8(5) where the non-labeled M13-HT primer was used as a substrate. Also, at the same time, incorporation activity was determined by the method described in Example 2(1) where an activated DNA was used as a substrate. The results are shown in Figure 11. The ratio of the primer extension activity to the incorporation activity for the two fractions was determined such that the ratio of 0.65 was obtained for the about 320 kilodalton fraction, and the ratio of 0.29 was obtained for the about 76 kilodalton fraction. Therefore, there was found that the primer extension activity of Pfu DNA polymerase is enhanced by the formation of a complex with F7.

Example 11

(1) Selection of Cosmid Clones Carrying Gene Encoding Homologs of RFC Large Subunit

[0143] Regarding the amino acid sequence of the RFC large subunit of *Methanococcus jannaschii* [Science, 273, 1058-1073 (1996)], homology to the amino acid sequences of PFU-RFC small subunits without carrying intein described in Example 9 was examined. In reference to the amino acid sequence of a region highly conserved among

them, the primer RFLS15 for searching the gene encoding the RFC large subunit was synthesized. The nucleotide sequence of the primer RFLS15 is shown in SEQ ID NO: 60 in Sequence Listing. PCR was carried out using a combination of this primer with the above primer RF-F1 corresponding to a similar amino acid sequence existing in the two subunit proteins of RFC with *Pyrococcus furiosus* genomic DNA as a template. The PCR was carried out using a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase, 0.25 μg of template DNA and 100 pmol each of primers. Of the two kinds of DNA fragments amplified by this PCR, an amplified DNA fragment of about 630 bp, of which size differs from the anticipated size of the amplification product derived from the PFU-RFC small subunit gene was isolated. This DNA fragment was subcloned into plasmid vector pUC119, and its nucleotide sequence was determined. Thereafter, in reference to the nucleotide sequence determined, the primers RFLS-S3 and RFLS-S4, of which nucleotide sequences are shown in SEQ ID NOs: 61 and 62 in Sequence Listing, were then synthesized.

[0144] PCR was carried out using these two primers with the cosmid DNA prepared in Example 1 as a template, whereby selecting cosmid clones assumed to carry the gene encoding homologs of the RFC large subunit. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme in 30 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (2 minutes). As a result, Cosmid Clone Nos. 254, 310, 313, 377 and 458 were found to carry the desired gene (PFU-RFCLS gene). These Cosmid Clone Numbers were identical to the above cosmid clones carrying the PFU-RFC gene. With this in mind, the nucleotide sequence of the DNA insert in the plasmid pRFS254NdB shown in SEQ ID NO: 55 in Sequence Listing was examined, and it was found that a homolog (PFU-RFCLS) of the RFC large subunit was encoded by the open reading frame starting at No. 3109 of the sequence immediately downstream of the PFU-RFC gene. However, this plasmid pRFS254NdB did not harbor a full length of the PFU-RFCLS gene.

(2) Subcloning of PFU-RFCLS Gene

[0145] In order to isolate a DNA fragment carrying the full length of the PFU-RFCLS gene, Clone No. 254 above was digested with *Nhe*I, and the various DNA fragments obtained were cut out, and each was subcloned into plasmid vector pTV118N (manufactured by Takara Shuzo Co., Ltd.). PCR was carried out using RFLS-S3 and RFLS-S4 as primers with each of the recombinant plasmids obtained as a template, in order to examine whether or not the PFU-RFCLS gene is present. As a result, an *Nhe*I fragment of about 11 kb was found to carry the RFLS gene. The plasmid resulting from insertion of this *Nhe*I fragment into pTV118N was named the plasmid pRFLSNh. In addition, a restriction endonuclease map of the DNA insert contained in this plasmid was prepared, and the results as shown in Figure 12 were obtained.

[0146] Furthermore, the nucleotide sequence of the DNA insert contained in this plasmid was determined by the dideoxy method. Of the nucleotide sequence determined, the nucleotide sequence of the open reading frame portion encoding PFU-RFCLS is shown in SEQ ID NO: 63 in Sequence Listing. The amino acid sequence of PFU-RFCLS deduced from the sequence is shown in SEQ ID NO: 64 in Sequence Listing.

Example 12

(1) Selection of Cosmid Clones Carrying F5 Gene

[0147] On the basis of the N-terminal amino acid sequence of F5 obtained in Example 3, the primers F5-1-1 and F5-2, of which nucleotide sequences are shown in SEQ ID NO: 65 and 66, respectively, in Sequence Listing, were synthesized. First PCR was carried out using 100 pmol each of F5-1-1 and the cassette primer C1 (manufactured by Takara Shuzo Co., Ltd.) with 1 μl of the *Pst*I cassette DNA prepared in Example 4 as a template. Second PCR was carried out using 100 pmol of both F5-2 and the cassette primer C2 (manufactured by Takara Shuzo Co., Ltd.) with 1 μl of the above reaction mixture as a template. This second PCR was carried out using TaKaRa PCR amplification kit (manufactured by Takara Shuzo Co., Ltd.) in accordance with the instructions attached. An amplified DNA fragment of about 900 bp was subcloned into plasmid vector pTV118N (manufactured by Takara Shuzo Co., Ltd.). The plasmid obtained was named pF5P2, and its nucleotide sequence was determined. Thereafter, on the basis of the sequence determined, primers F5S1 and F5S2, of which nucleotide sequences are shown in SEQ ID NOs: 67 and 68, respectively, in Sequence Listing, were synthesized. PCR was carried out using these F5S1 and F5S2 with the cosmid DNA described in Example 1 as a template, whereby selecting cosmid clones carrying the F5 gene. This PCR was carried out using the TaKaRa PCR amplification kit in accordance with the instructions attached. As a result, there were found that Cosmid Clone Nos. 15, 96, 114, 167, 277, 348, 386, 400, 419, 456, 457 and 484 carry the F5 gene. These Cosmid Clone Numbers were identical to the cosmid clones carrying the F7 gene. With this in mind, the nucleotide sequence shown in SEQ ID NO: 41 in Sequence Listing was examined, and it was found that a portion on or after No. 892, which is downstream of the F7 gene on the sequence, carries a first half of the F5 gene.

(2) Subcloning of F5 Gene

[0148] In order to subclone the F5 gene, a restriction endonuclease map for *Nco*I, *Bam*HI, *Pst*I, *Hind*III and *Nde*I (manufactured by Takara Shuzo Co., Ltd.) in the neighborhood of the F5 gene was prepared using the plasmid pF7-HH-18 obtained in Example 8 and the above plasmid pF5P2, and the results as shown in Figure 13 were obtained.

[0149] On the basis of the restriction endonuclease map shown in Figure 13, Cosmid Clone No. 15 was digested with *Nde*I, and a fragment of about 900 bp was cut out and subcloned into plasmid vector pTV118Nd. As to the recombinant plasmid obtained, a plasmid resulting from insertion of the F5 gene in the orthodox orientation with respect to the *lac* promoter was named pF5NNF-1.

(3) Determination of Nucleotide Sequence of DNA Fragment Carrying F5 Gene

[0150] The nucleotide sequence of the DNA insert in the above plasmid pF5NNF-1 was determined by the dideoxy method. As a result of analyzing the nucleotide sequence determined, there was found an open reading frame encoding a protein of which N-terminal amino acid sequence is identical to that of F5. The nucleotide sequence of this open reading frame is shown in SEQ ID NO: 69 in Sequence Listing, and the amino acid sequence of F5 as deduced from the above nucleotide sequence is shown in SEQ ID NO: 70 in Sequence Listing. This amino acid sequence was searched for homology to the amino acid sequences of known proteins, and as a result, proteins homologous thereto were not found.

(4) Construction of Plasmid for F5 Expression

[0151] PCR was carried out using the primers F5Nco and F5CBam, of which nucleotide sequences are shown in SEQ ID NOs: 71 and 72, respectively, in Sequence Listing, with the above plasmid pF5NNF-1 as a template. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase. Using 1 ng of a template DNA and 20 pmol each of both of the primers, the reaction was carried out in 25 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (2 minutes). An amplified DNA fragment of an about 640 base pairs was digested with *Nco*I and *Bam*HI (both manufactured by Takara Shuzo Co., Ltd.), and the fragment obtained was ligated with pET15b (manufactured by Novagen), previously linearized with *Nco*I and *Bam*HI. This plasmid was named pF5NBPET. Of the DNA insert in the plasmid, the region amplified by PCR was analyzed by the dideoxy method to determine its nucleotide sequence. There was confirmed that there is no mutation caused by PCR.

[0152] *Escherichia coli* HMS174(DE3)/pF5NBPET, *Escherichia coli* HMS174(DE3) transformed with the plasmid pF5NBPET, was evaluated for F5 expression, and there was demonstrated that a protein of a molecular weight corresponding to F5 in the culture of the transformant is expressed.

<u>Example 13</u>

(1) Subcloning of F3 Gene

[0153] On the basis of the N-terminal amino acid sequence of F3 obtained in Example 3, the primers F3-1 and F3-3-1, of which nucleotide sequences are shown in SEQ ID NOs: 73 and 74 in Sequence Listing, were synthesized. First PCR was carried out using 100 pmol of the primer F3-1 and 20 pmol of the cassette primer C1 with 1 µl of the *Bgl*II/*Sau*3AI cassette DNA of Example 4 as a template. With 1 µl of the above reaction mixture as a template, second PCR was carried out using F3-3-1 and the cassette primer C2. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme in 30 cycles for the first PCR and 25 cycles for the second, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 45°C (30 seconds) - 72°C (2 minutes). An amplified DNA fragment of about 500 bp by this reaction was subcloned into plasmid vector pTV118N, and a part of its nucleotide sequence was determined by the dideoxy method using M4 and RV primers (manufactured by Takara Shuzo Co., Ltd.). On the basis of the sequence determined, the primers F3S1, F3S2, F3S3 and F3S4, of which nucleotide sequences are shown in SEQ ID NOs: 75, 76, 77 and 78 in Sequence Listing, were then synthesized. PCR was carried out using these F3S1 and F3S2 primers with the cosmid DNA prepared in Example 1 as a template, and cosmid clones carrying the F3 gene were searched. As a result, there was found no cosmid clone assumed to carry the F3 gene. With this in mind, PCR was carried out using the primer F3S3 or F3S4 and the primer C2 with each cassette DNA of Example 4 as a template. As a result of mapping of the restriction endonuclease recognition sites in the neighborhood of the F3 gene, there was anticipated that the F3 gene is present in a fragment of about 2.6 kb between the *Sal*I site and the *Hind*III site. On the basis of the results, 4 µg of *Pyrococcus furiosus* genomic DNA was digested with *Sal*I and *Hind*III, and thereafter a DNA fragment of about 2.6 kb was collected and subcloned into

pTV118N vector. PCR was carried out using the primer F3S4 and the primer RV-N (manufactured by Takara Shuzo Co., Ltd.) with each of the recombinant plasmids thus obtained as a template, to examine for the presence of the F3 gene. As a result, a plasmid harboring a 2.6 kb *Sal*I-*Hind*III fragment carrying the F3 gene was obtained, and this plasmid was named the plasmid pF3SH92. *Escherichia coli* JM109/pF3SH92, *Escherichia coli* JM109 transformed with this plasmid, was examined for F3 expression, and as a result, there was confirmed that a protein having a molecular weight corresponding to F3 is expressed.

(2) Determination of Nucleotide Sequence of DNA Fragment Carrying F3 Gene

[0154] The nucleotide sequence of the DNA insert in the above plasmid pF3SH92 was determined by the dideoxy method. As a result of analyzing the nucleotide sequence determined, there was found an open reading frame encoding a protein of which N-terminal amino acid sequence is identical to that of F3. The nucleotide sequence of this open reading frame is shown in SEQ ID NO: 79 in Sequence Listing, and the amino acid sequence of F3 as deduced from the nucleotide sequence is shown SEQ ID NO: 80, respectively, in Sequence Listing. This amino acid sequence was searched for homology to the amino acid sequences of known proteins, and as a result, the amino acid sequence is found to be homologous to *Mycoplana ramosa*-derived acetyl polyamine aminohydrase [*Journal of Bacteriology*, **178**, 5781-5786 (1996)] and human histone deacetylase [Science, **272**, 408-411 (1996)].

Example 14

[0155] In the following Example, the activities of commercially available enzymes are shown on the basis of the labeling for individual enzymes. Also, reaction mixtures containing commercially available enzymes were prepared in accordance with the manuals for the respective enzymes, or using the reaction buffers attached thereto, unless otherwise specified. PCR was carried out using GeneAmp PCR System 9600 (manufactured by Perkin-Elmer).

(1) Preparation of Anti-PFU-RFC Antibody

[0156] The PFU-RFC authentic sample of Example 9(6) was diluted so as to have a concentration of 1 mg/100 µl with 50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol and 75 mM NaCl, and the mixture was emulsified with an equal volume of Freund's complete adjuvant. This emulsion was subcutaneously injected to rabbits at 50 µl per injection four times at 3-week intervals. Ten days after final immunization, whole blood was extracted. After allowing to stand at room temperature for 60 minutes, the extracted blood was centrifuged to yield 50 ml of an antiserum containing the anti-PFU-RFC polyclonal antibody. To 20 ml of this antiserum, 20 ml of a saturated solution of ammonium sulfate was added, and the mixture was gently stirred at 4°C for 45 minutes and subsequently centrifuged. The precipitate obtained was suspended in 5 ml of 20 mM sodium phosphate buffer, pH 7.0, and thrice subjected to 2-hour dialysis against 2 liters of the same buffer as a dialysate. After dialysis, 14 ml of the solution was applied to Protein A column (manufactured by Pharmacia), previously equilibrated with 20 mM sodium phosphate buffer (pH 7.0). After the column was washed with the same buffer, the elution was carried out with 0.1 M sodium citrate buffer (pH 3.0). After the anti-PFU-RFC antibody eluted was neutralized with 1 M Tris-HCl, pH 9.0, the mixture was then concentrated using Centriflow CF-50 and subjected to exchange with coupling buffer (0.5 M NaCl, 0.2 M $NaHCO_3$, pH 8.3) using PD-10 column to prepare a solution containing the anti-PFU-RFC antibody.

(2) Preparation of Anti-PFU-RFC Antibody Column

[0157] HiTrap NHS-activated column (manufactured by Pharmacia) was washed with 6 ml of 1 mM HCl, and thereafter 0.95 ml of the above anti-PFU-RFC polyclonal antibody solution (containing 3.8 mg equivalent of the anti-PFU-RFC antibody) was applied thereto. Subsequently, an anti-PFU-RFC antibody column was prepared in the same manner as Example 2(3).

(3) Purification of Complex Containing PFU-RFC Using Anti-PFU-RFC Antibody Column

[0158] *Pyrococcus furiosus* DSM3638 was cultured in the same manner as the method described in Example 1 to yield cells in 10 liters of culture medium. These cells were suspended in 33 ml of buffer C (50 mM Tris-HCl, pH 8.0, 0.1 mM ATP) containing 2 mM PMSF, and the suspension was treated with an ultrasonic disrupter. The disrupted solution was centrifuged at 12,000 rpm for 10 minutes, and 38 ml of the supernatant obtained was applied to the anti-PFU-RFC antibody column, previously equilibrated with buffer C containing 0.1 M NaCl. After washing with buffer C containing 0.1 M NaCl, the column was heated at 85°C for 1 hour, and the PFU-RFC complex was eluted with buffer C containing 0.1 M NaCl. This eluate was subjected to SDS-PAGE (12.5% polyacrylamide gel; 25 mM Tris-HCl, 192 mM glycine, 0.1%

SDS, pH 8.4 used as electrophoresis buffer). The gel after electrophoresis was stained with Coomassie brilliant blue R-250 by a conventional method, and as a result, in addition to the band of PFU-RFC, one band at a position for 33 kilodaltons, which corresponds to the above F7, and two bands near 60 kilodaltons were detected.

[0159] With this in mind, the N-terminal amino acid sequences of the proteins existing in these three bands were analyzed in the same manner as Example 3(2). As a result, as shown in Figure 14, the N-terminal amino acid sequence of the protein at a position corresponding to the above F7 was found to be identical to that of F7, and each of the N-terminal amino acid sequences of the two kinds of proteins near 60 kilodaltons was found to be identical to the above N-terminal amino acid sequence of the PFU-RFCLS.

[0160] Next, the amounts of the PFU-RFC, PFU-RFCLS and F7 proteins in this eluate were quantified by the amount of Coomassie brilliant blue bound thereto. The eluate was subjected to SDS-PAGE (12.5% polyacrylamide gel; 25 mM Tris-HCl, 192 mM glycine, 0.1% SDS, pH 8.4 used as electrophoresis buffer). The gel after electrophoresis was stained with Coomassie brilliant blue R-250 by a conventional method, and thereafter the band was cut out and treated with 500 μl of 70% formic acid to extract the Coomassie brilliant blue, and the absorbance at 630 nm was determined. On the basis of a calibration curve prepared using the F7 authentic sample of Example 8(4) and the PFU-RFC authentic sample of Example 9(6), each of a known concentration, it was found that 208 μg of PFU-RFC, 55 μg of PFU-RFCLS and 51 μg of the F7 protein were contained in 500 μl of the eluate. The complex constituted by the three proteins PFU-RFC, PFU-RFCLS and F7 as described above is hereinafter referred to as RFC-N complex.

(4) Effects of RFC-N Complex on Primer Extension Reactions

[0161] In order to examine the effects of the RFC-N complex obtained in Example 14(3) on the primer extension reactions of various polymerases, the activities of Pfu polymerase C and Pfu DNA polymerase (α-type DNA polymerase, manufactured by STRATAGENE) were compared between cases where the RFC-N complex was added and cases where only its constituent F7 was added. The DNA polymerase activities were determined in the same manner as the method described in Example 8(5), except that 50 fmol of Pfu polymerase C or Pfu DNA polymerase was used. For the determination of the DNA polymerase activities, one prepared by annealing the HT primer, which is a synthetic oligonucleotide of 45 bases, to M13 phage single-stranded DNA (M13mp18ss DNA, manufactured by Takara Shuzo Co., Ltd.), was used as shown in Example 8(5) (M13-HT primer). The nucleotide sequence of the HT primer is shown in SEQ ID NO: 42 in Sequence Listing. The results for Pfu DNA polymerase are shown in Figure 15. The amounts of F7 and the RFC-N complex added are expressed in the molar numbers of F7 and RFC-N complex contained in the reaction mixture. As shown in Figure 15, the RFC-N complex showed higher increase in the activity to Pfu DNA polymerase than that of F7 alone.

[0162] Furthermore, the primer extension activity was studied by the method described in Example 8(5). Reaction mixtures for determination were prepared with the following compositions: 1) 100 fmol of F7, 2) 0.05 μl of the RFC-N complex (containing 60 fmol of F7), 3) 10 fmol of Pfu polymerase C, 4) 10 fmol of Pfu polymerase C + 100 fmol of F7, 5) 100 fmol of Pfu polymerase C + 0.05 μl of the RFC-N complex, 6) 20 fmol of F7, 7) 0.02 μl of the RFC-N complex (containing 24 fmol of F7), 8) 10 fmol of Pfu DNA polymerase, 9) 10 fmol of Pfu DNA polymerase + 20 fmol of F7, 10) 10 fmol of Pfu DNA polymerase + 0.02 μl of the RFC-N complex. To 1 μl of each reaction mixture for determination, 9 μl of a reaction mixture [20 mM Tris-HCl (pH 9.0), 15 mM $MgCl_2$, 2 mM 2-mercaptoethanol, 40 μM each of dATP, dGTP, dCTP and dTTP] containing 0.01 μg/μl ^{32}P-labeled M13-HT primer was added, and the reaction was carried out at 75°C for 2.5 minutes. After termination of the reaction, the reaction mixture was cooled with ice to stop the reaction, and 1 μl of 200 mM EDTA and 5 μl of a reaction stopper (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) were further added thereto, and the mixture was subjected to thermal denaturation treatment at 95°C for 5 minutes. After 1.6 μl of this reaction mixture was electrophoresed using 6% polyacrylamide gel containing 8 M urea, an autoradiogram was prepared.

[0163] Next, in order to analyze primer extension reaction products of longer chains, the analysis was carried out by the method described in Example 8(5). To 1 μl of each of sample solutions 1) to 10) above, 9 μl of a reaction mixture [20 mM Tris-HCl, pH 9.0, 15 mM $MgCl_2$, 2 mM 2-mercaptoethanol, 40 μM each of dATP, dGTP, dCTP and dTTP, 84 nM [α-^{32}P]-dCTP] containing M13-HT primer to have a final concentration of 0.01 μg/μl was added, and the mixture was reacted at 75°C for 2.5 minutes. After termination of the reaction, to the ice cooled reaction mixture, 1.11 μl of 200 mM EDTA, 1.23 μl of 500 mM NaOH, and 2.47 μl of 6-fold concentrated loading buffer (0.125% bromophenol blue, 0.125% xylene cyanol, 9% glycerol) were sequentially added. After 6 μl of this mixture was electrophoresed using 0.5% alkaline agarose gel, an autoradiogram was prepared.

[0164] In either case of Pfu polymerase C and Pfu DNA polymerase, the amount of long-chain extension products increased in the case where the RFC-N complex was added as compared to the case of F7 alone.

[0165] The chain lengths of the long-chain extension products were found to be up to about 7.2 kb, a full length of the template, in either of the polymerases used, in the case of F7 alone and of the RFC-N complex.

Example 15 Construction of Plasmid for rRFC-M Expression

[0166]

(1) A plasmid for simultaneously expressing PFU-RFCLS and PFU-RFC was constructed. In reference to the nucleotide sequence determined in Example 11(2), the primer RFLS-NdeN, of which nucleotide sequence is shown in SEQ ID NO: 81 in Sequence Listing, and RFLS-S9, of which nucleotide sequence is shown in SEQ ID NO: 82, were synthesized. PCR was carried out using both of these primers with the above plasmid pRFLSNh as a template. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme, 10 ng of the plasmid pRFLSNh and 20 pmol each of the primers in 30 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (3 minutes). An *Nde*I-*Pst*I fragment of about 920 bp isolated after digesting an amplified DNA fragment obtained by PCR with *Nde*I and *Pst*I, a *Pst*I-*Eco*RI fragment of about 600 bp isolated from the plasmid pRFLSNh obtained in Example 11(2), and an *Eco*RI-*Bam*HI fragment of about 2 kb isolated from the plasmid pRFS254SNc obtained in Example 9(4) were mixed and subcloned between the *Nde*I and *Bam*HI sites of plasmid vector pTV119Nd. The recombinant plasmid thus obtained was named pRFC10. In addition, *Escherichia coli* JM109 transformed with the plasmid was named *Escherichia coli* JM109/pRFC10. This transformant was found to possess a high level of expression of PFU-RFCLS and PFU-RFC.

(2) Determination of Nucleotide Sequence of Genes Encoding PFU-RFCLS and PFU-RFC

Of the DNA insert in the plasmid pRFC10 obtained in Example 15(1), the region amplified by PCR was analyzed by the dideoxy method to determine its nucleotide sequence, and it was confirmed that there is no mutation caused by PCR. From this result and the results of Example 9(3) and Example 11(2), the nucleotide sequence of the gene encoding PFU-RFCLS and PFU-RFC without carrying intein was determined. The nucleotide sequence of the genes encoding PFU-RFCLS and PFU-RFC without carrying intein thus obtained is shown in SEQ ID NO: 83 in Sequence Listing, and its restriction endonuclease map is shown in Figure 16.

Example 16 Preparation of rRFC-M Authentic Sample

[0167] *Escherichia coli* JM109/pRFC10 obtained in Example 15(1) was cultured for 16 hours in 500 ml x 4 of LB medium (10 g/l trypton, 5 g/l yeast extract, 5 g/l NaCl, pH 7.2), in which ampicillin was present at a concentration of 100 µg/ml, and IPTG is present at 1 mM. After harvesting, cells were suspended in 35.9 ml of sonication buffer [50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol, 10% glycerol, 2 mM PMSF (phenylmethanesulfonyl fluoride)], and the suspension was treated with an ultrasonic disrupter. After centrifugation at 12,000 rpm for 10 minutes, a heat treatment was carried out at 80°C for 15 minutes. Thereafter, centrifugation at 12,000 rpm for 10 minutes was again carried out to yield 33.0 ml of a heat-treated enzyme solution. This solution was then applied to RESOUCE Q column (manufactured by Pharmacia), previously equilibrated with buffer A (50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol, 10% glycerol), and the applied solution was chromatographed using FPLC system (manufactured by Pharmacia). The elution was carried out on a linear concentration gradient of 0 to 500 mM NaCl.

[0168] As a result of analyzing the eluate by SDS-PAGE (12.5% polyacrylamide gel; 25 mM Tris-HCl, 192 mM glycine, 0.1% SDS, pH 8.4 used as electrophoresis buffer), PFU-RFCLS and PFU-RFC were both eluted at an NaCl concentration of 240 mM. When the eluate obtained from cells in which PFU-RFC was expressed alone as described in Example 9(6) was applied to RESOURCE Q column, the eluate was not adsorbed to RESOURCE Q column. On the other hand, when the eluate obtained from cells in which PFU-RFCLS and PFU-RFC were simultaneously expressed was applied to RESOURCE Q column, the eluate was adsorbed thereto, and PFU-RFCLS and PFU-RFC were simultaneously eluted at an NaCl concentration of 240 mM, as described above. From the results, it was demonstrated that these two proteins have formed a complex. This complex is hereinafter referred to as rRFC-M complex.

[0169] After 4.8 ml of an enzyme solution obtained by collecting the rRFC-M complex fraction was concentrated using Centriflow CF50, the concentrate was subjected to exchange with buffer A containing 150 mM NaCl using PD-10 column (manufactured by Pharmacia), and 3.5 ml of the solution was applied to Heparin column (manufactured by Pharmacia), previously equilibrated with buffer A containing 150 mM NaCl. Using FPLC system, the chromatogram was developed on a linear concentration gradient from 150 mM to 650 mM NaCl, and an rRFC-M complex fraction eluted at 450 mM NaCl was obtained. Using Centricon-10 (manufactured by Amicon), 3.9 ml of this fraction was concentrated, and 115 µl of the concentrate was applied to Superdex 200 gel filtration column (manufactured by Pharmacia), previously equilibrated with 50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol and 75 mM NaCl. The elution was carried out with the same buffer, and the rRFC-M complex was found to have a retention time of 26.3 minutes. From the comparative results with the position of the elution of a molecular weight marker under the same conditions, the molecular weight of the rRFC-M complex was calculated as about 370 kilodaltons.

[0170] Furthermore, in order to determine the compositional ratio of each unit in the rRFC-M complex, the above eluted fraction of a molecular weight of about 370 kDa was subjected to SDS-PAGE.

[0171] The gel after electrophoresis was stained with Coomassie brilliant blue R-250 by a conventional method, and thereafter the bands of the PFU-RFCLS and PFU-RFC proteins were cut out and extracted with 500 µl of 70% formic acid. The absorbance at 630 nm of each extract was determined, and the results were compared with the calibration curve prepared by using PFU-RFC prepared in Example 9(6), and whereby the amount of each protein was determined and the molar number was calculated.

[0172] As a result, PFU-RFCLS and PFU-RFC were found to exist in a 1:4 ratio. Based on the fact that the molecular weight of the rRFC-M complex as calculated by the gel filtration described above was about 370 kDa, the rRFC-M complex was assumed to be formed by two molecules of PFU-RFCLS and eight molecules of PFU-RFC. With this in mind, the molar number was calculated, taking the above rRFC-M complex as 1 unit.

Example 17 Construction of Plasmid F3 Expression

[0173]

(1) PCR was carried out using the primer F3Nd, of which nucleotide sequence is shown in SEQ ID NO: 84 in Sequence Listing, and the F3S2 primer, of which nucleotide sequence is shown in SEQ ID NO: 76, with the plasmid pF3SH92 as prepared in Example 13 as a template. The PCR was carried out in a reaction mixture of the same composition as that used in Example 5(1) using Pfu DNA polymerase as an enzyme, 1 ng of the plasmid pF3SH92 and 20 pmol each of the primers in 30 cycles, wherein one cycle comprises a process consisting of at 94°C (30 seconds) - 55°C (30 seconds) - 72°C (1 minute). An *Nde*I-*Pst*I fragment of about 0.5 kb isolated after digestion of an amplified DNA fragment obtained by PCR with *Nde*I and *Pst*I, and a *Pst*I-*Eco*RI fragment of about 1.1 kb isolated from the plasmid pF3SH92 were mixed and subcloned between the *Nde*I and *Eco*RI sites of plasmid vector pTV119Nd. The recombinant plasmid thus obtained was named pF3-19. In addition, *Escherichia coli* JM109 transformed with the plasmid was named *Escherichia coli* JM109/pF3-19. The transformant was found to possess high expression of F3.

(2) Determination of Nucleotide Sequence of Gene Encoding F3

Of the DNA insert in the plasmid pF3-19, obtained in Example 17(1), the region amplified by PCR was analyzed by the dideoxy method to determine its nucleotide sequence, and confirmed that there is no mutation caused by PCR.

Example 18 Preparation of Purified F3 Authentic Sample

[0174] *Escherichia coli* JM109/pF3-19 obtained in Example 17(1) was cultured for 16 hours in 500 ml x 4 of LB medium (10 g/liter trypton, 5 g/liter yeast extract, 5 g/liter NaCl, pH 7.2) in which ampicillin was present at a concentration of 100 µg/ml. After harvesting, cells were suspended in 50 ml of sonication buffer [50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol, 10% glycerol, 2 mM PMSF (phenylmethanesulfonyl fluoride)], and the suspension was treated with an ultrasonic disrupter. After centrifugation at 12,000 rpm for 10 minutes, the supernatant was subjected to heat treatment at 80°C for 15 minutes. Thereafter, centrifugation at 12,000 rpm for 10 minutes was again carried out to yield a heat-treated supernatant. Forty-four milliliters of the heat-treated supernatant was applied to RESOURCE Q column (manufactured by Pharmacia), previously equilibrated with buffer A described in Example 16, and the applied solution was chromatographed using FPLC system (manufactured by Pharmacia). The chromatogram was developed on a linear concentration gradient from 0 to 500 mM NaCl. To 11 ml of a solution of the fraction containing F3 eluted at 140 mM to 240 mM NaCl, 5.5 ml of buffer A containing 3 M ammonium sulfate was added, and this solution was applied to HiTrap butyl column (manufactured by Pharmacia), previously equilibrated with buffer A containing 1 M ammonium sulfate. After the column was washed with buffer A containing 1 M ammonium sulfate using FPLC system, F3 was eluted with buffer A containing 0.5 M ammonium sulfate. Six milliliters of this fraction was applied to HiTrap phenyl column (manufactured by Pharmacia), previously equilibrated with buffer A containing 0.5 M ammonium sulfate. After the column was washed with buffer A containing 0.5 M ammonium sulfate using FPLC system, F3 was eluted with buffer A. Using Centricon-10 (manufactured by Amicon), 9.5 ml of this fraction was concentrated, and 155 µl of the concentrate was applied to Superdex 200 gel filtration column (manufactured by Pharmacia), previously equilibrated with 50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol and 75 mM NaCl. The elution was carried out with the same buffer, and as a result, F3 was eluted at a position corresponding to a retention time of 42.1 minutes. From the comparative results in the position of the elution of a molecular weight marker under the same conditions, a molecular weight of about 25 kilodaltons was anticipated. On the basis that the theoretical value of the molecular weight of F3 is 37 kilodaltons, F3 is deduced to be a monomer.

Example 19 Preparation of Purified F5 Authentic Sample

[0175] *Escherichia coli* HMS174(DE3)/pF5NBPET, *Escherichia coli* HMS174(DE3) transformed with the plasmid pF5NBPET obtained in Example 12(4), was cultured for 16 hours in 500 ml x 4 of LB medium (10 g/liter trypton, 5 g/liter yeast extract, 5 g/liter NaCl, pH 7.2) in which ampicillin was present at a concentration of 100 μg/ml. After harvesting, cells were suspended in 61 ml of sonication buffer, and the suspension was treated with using an ultrasonic disrupter. The disrupted cells were centrifuged at 12,000 rpm for 10 minutes, and thereafter the supernatant was subjected to heat treatment at 80°C for 15 minutes. Thereafter, centrifugation at 12,000 rpm for 10 minutes was again carried out to yield a heat-treated supernatant. To 60.5 ml ammonium sulfate, 8.71 g of ammonium sulfate was added, and the mixture was stirred at 4°C for 2 hours, and thereafter centrifugation at 12,000 rpm for 10 minutes was carried out. The precipitate was dissolved in 19 ml of buffer A and dialyzed against buffer A. The enzyme solution after dialysis was applied to RESOURCE Q column (manufactured by Pharmacia), previously equilibrated with buffer A, and the applied solution was chromatographed using FPLC system (manufactured by Pharmacia). The chromatogram was developed on a linear concentration gradient from 0 to 500 mM NaCl. Using Centricon-10 (manufactured by Amicon), 11 ml of a solution of a fraction containing F5 eluted at 350 mM to 450 mM NaCl was concentrated, and 222 μl of the concentrate was applied to Superdex 200 gel filtration column (manufactured by Pharmacia), previously equilibrated with 50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol and 75 mM NaCl. The elution was carried out with the same buffer, and as a result, F5 was eluted at a position corresponding to a retention time of 32.5 minutes. From the comparative results with the position of the elution of a molecular weight marker under the same conditions, a molecular weight of about 145 kilodaltons was anticipated. This molecular weight corresponds to the case where F5 has formed a heptamer.

Example 20 Preparation of Primers

[0176] On the basis of the nucleotide sequence of λDNA, eight kinds of primers, i.e., λ1B to λ5 and λ7 to λ9, were synthesized. The nucleotide sequences of the primers λ1B to λ5 and λ7 to λ9 are shown in SEQ ID NOs: 85 to 92, respectively, in Sequence Listing. The chain lengths of DNA fragments amplified by PCR using combinations of these primers with λDNA as a template are shown in Table 4.

Table 4

Primer Pairs	Chain Length of DNA Fragment Amplified
λ1B / λ2	0.5 kb
λ1B / λ3	1 kb
λ1B / λ4	2 kb
λ1B / λ5	4 kb
λ1B / λ7	8 kb
λ1B / λ8	10 kb
λ1B / λ9	12 kb

Example 21 Effects of F1 Protein on DNA Polymerase

[0177] The effects of the F1 protein obtained in Example 5 on PCR were examined. In order to carry out an amplification reaction of 1 to 4 kb DNA fragments using λDNA as a template, each of the primers λ1B and λ3, the primers λ1B and λ4, and the primers λ1B and λ5, were used as primer pairs to prepare reaction mixtures of the compositions shown below: 10 mM Tris-HCl, pH 9.2, 75 mM KCl, 6 mM $MgCl_2$, 0.4 mM each of dATP, dCTP, dGTP and dTTP, 0.01% BSA and 1.25 units of Pfu polymerase C, 500 pg of template DNA, 5 pmol each of the primers, 173 pmol of the F1 protein (final volume being 25 μl). Using each reaction mixture, the reaction was carried out in 30 cycles, wherein one cycle comprises a process consisting of at 98°C, 0 second - 68°C, 0 second. The phrases "98°C, 0 second", "68°C, 0 second" etc. as used in the present specification indicate that the reaction apparatus was programmed so that the setting temperature is immediately shifted to the next one when the setting temperature is reached.

[0178] After termination of the reaction, 5 μl of the reaction mixture was electrophoresed on 1% agarose gel (manufactured by Takara Shuzo Co., Ltd.) to confirm amplified fragments.

[0179] As a result, the amplification of 1 kb, 2 kb and 4 kb DNA fragments, depending on the primer pairs used, was

confirmed. On the other hand, when the above reaction mixture without the addition of the F1 protein was subjected to PCR under the above reaction conditions, no amplified fragments could be confirmed.

Example 22 Effects of F1, F3 and F5 Proteins on DNA Polymerase

[0180] The effects of the F1 protein obtained in Example 5, the F3 protein obtained in Example 18 and the F5 protein obtained in Example 19 were used to investigate the amplification of a 6 kb DNA fragment by PCR with λDNA as a template. Reaction mixtures of the same compositions as those used in Example 21 were prepared, except that the primers λ1 and λ6 were used as a primer pair. The F1 protein was added in an amount of 173 pmol, the F3 protein was added in an amount of 10 pmol, and the F5 protein was added in an amount of 1 pmol, respectively, to make up a final volume of 25 µl. Using each reaction mixture, the reaction was carried out in 30 cycles, wherein one cycle comprises a process consisting of at 98°C, 1 second - 68°C, 2 minutes. After termination of the reaction, 5 µl of the reaction mixture was electrophoresed on 1% agarose gel to confirm amplified fragments.

[0181] As a result, the amplification of a 6 kb DNA fragment was confirmed in the presence of any of the F1, F3 and F5 proteins. On the other hand, when these proteins were not added, no amplified fragments could not be confirmed.

Example 23 Effects of F2 and F4 Proteins on DNA Polymerase

[0182] The effects of the F2 protein obtained in Example 6 and the F4 protein obtained in Example 7 were used to investigate the amplification reaction of a 4 kb DNA fragment by PCR with λDNA as a template. Reaction mixtures of the same compositions as those used in Example 21 were prepared, except that the primers λ1B and λ5, as a primer pair, 0.75 units of Pfu polymerase C and 1 ng of template λDNA were used. The F2 protein and the F4 protein were each added in an amount of 1.095 pmol to the reaction mixture to make up a final volume of 25 µl. Using each reaction mixture, the reaction was carried out in 25 cycles, wherein one cycle comprises a process consisting of at 94°C, 30 seconds - 55°C, 30 seconds - 72°C, 2 minutes. After termination of the reaction, 5 µl of the reaction mixture was electrophoresed on 1% agarose gel to confirm amplified fragments.

[0183] As a result, the amplification of a 4 kb fragment was confirmed in the presence of any of the F2 and F4 proteins. On the other hand, when these proteins were not added, no amplified fragment was confirmed.

Example 24 Effects of rRFC-M Complex on DNA Polymerases

[0184] In order to examine the effects of the rRFC-M complex on the primer extension reactions of various polymerases, the activities of Pfu polymerase C and Pfu DNA polymerase (α-type DNA polymerase, manufactured by STRATAGENE) were compared for cases where the rRFC-M complex and F7 are coexistent, and for cases where F7 exists alone.

[0185] DNA polymerase activities were determined in the same manner as the method described in Example 8(5), except that 50 fmol of Pfu polymerase C or Pfu DNA polymerase was used, and that 400 fmol of the rRFC-M complex and 0 to 200 fmol of F7 were added. The results of the case of using Pfu DNA polymerase are shown in Figure 17. The effects on Pfu DNA polymerase were such that the activity was more elevated in the case of coexistence of the rRFC-M complex and F7 than the case of F7 alone. In addition, the effects on Pfu polymerase C showed the same tendency as those of Pfu DNA polymerase.

Example 25 Effects of Coexistence of rRFC-M Complex and F7 Protein on PCR

[0186] In order to carry out an amplification reaction of a 4 kb DNA fragment using λDNA as a template, reaction mixtures of the same compositions as those used in Example 21 were prepared, except that the primers λ1B and λ5 and 0.375 units of Pfu polymerase C were used. The rRFC-M complex was added in an amount of 312.5 fmol, and the F7 protein was added in an amount of 125 fmol, respectively, to the reaction mixture to make up a final volume of 25 µl. Using each reaction mixture, the reaction was carried out in 30 cycles, wherein one cycle comprises a process consisting of at 98°C, 0 second - 68°C, 10 seconds. After termination of the reaction, 5 µl of the reaction mixture was electrophoresed on 1% agarose gel (manufactured by Takara Shuzo Co., Ltd.) to confirm amplified fragments.

[0187] As a result, the amplification of a 4 kb DNA fragment, depending on the primer pair used, was confirmed in the case of the system where the rRFC-M complex and the F7 protein were coexistent. On the other hand, when these proteins were not added, no amplified fragments could be confirmed.

[0188] Furthermore, a similar experiment was carried out for an amplification reaction of 8 to 12 kb DNA fragments using λDNA as a template. Reaction mixtures of the same compositions as those used in Example 21 were prepared, except that each of the primers λ1B and λ7, the primers λ1B and λ8, and the primers λ1B and λ9 were used as primer pairs, and further 0.375 units of Pfu polymerase C, and 2.5 ng of template λDNA were used. The rRFC-M complex was

added in an amount of 312.5 fmol, and the F7 protein was added in an amount of 125 fmol, respectively, to the reaction mixture to make up a final volume of 25 μl. Using each reaction mixture, the reaction was carried out in 30 cycles, wherein one cycle comprises a process consisting of at 98°C, 0 second - 68°C, 3 minutes. After termination of the reaction, 5 μl of the reaction mixture was electrophoresed on 1% agarose gel (manufactured by Takara Shuzo Co., Ltd.) to confirm amplified fragments.

[0189] As a result, the amplification of 8 kb, 10 kb and 12 kb DNA fragments, depending on the primer pairs used, was confirmed in the case of the system where the rRFC-M complex and the F7 protein were coexistent. On the other hand, when these proteins were not added, only a 8 kb DNA fragment was confirmed.

Example 26 Effects of Coexistence of rRFC-M Complex and F7 Protein on Pfu DNA Polymerase

[0190] In order to carry out an amplification reaction of a 4 kb DNA fragment using λDNA as a template, using each of the primers λ1B and λ3, the primers λ1B and λ4, and the primers λ1B and λ5, as primer pairs, reaction mixtures of the compositions shown below were prepared: buffer supplied with Pfu DNA polymerase, 0.2 mM each of dATP, dCTP, dGTP and dTTP, and 0.5 units each of Pfu polymerase, 500 pg of template DNA, 2.5 pmol of each primer, 2.5 pmol of the rRFC-M complex protein, and 0.5 pmol of the F7 protein (final volume being 25 μl). Using each reaction mixture, the reaction was carried out in 25 cycles, wherein one cycle comprises a process consisting of at 94°C, 30 seconds - 55°C, 30 seconds - 72°C, 1 minute. After termination of the reaction, 5 μl of the reaction mixture was electrophoresed on 1% agarose gel to confirm amplified fragments.

[0191] As a result, the amplification of 1 kb, 2 kb and 4 kb DNA fragments, depending on the primer pairs used, was confirmed in the case of the system where the rRFC-M complex and the F7 protein were coexistent. On the other hand, when these proteins were not added, only 1 kb to 2 kb DNA fragments were confirmed.

Example 27 Effects of Coexistence of rRFC-M Complex and F7 Protein on Mixed DNA Polymerase

[0192] The effects of the coexistence of the rRFC-M complex and the F7 protein on PCR using a mixture of two kinds of DNA polymerases were examined.

[0193] In order to carry out an amplification reaction of a 1 kb DNA fragment using λDNA as a template, using the primers λ1B and λ3 as a primer pair, reaction mixtures of the compositions shown below were prepared: buffer supplied with TaKaRa LA Taq (Mg Plus), 0.4 mM each of dATP, dCTP, dGTP and dTTP, 1.25 units of LA Taq DNA polymerase (manufactured by Takara Shuzo Co., Ltd.), 500 pg of template DNA, 5 pmol of each primer, 62.5 fmol of the RFC complex protein, and 12.5 fmol of the F7 protein (final volume being 25 μl). Using each reaction mixture, the reaction was carried out in 30 cycles, wherein one cycle comprises a process consisting of at 98°C, 0 second - 68°C, 10 seconds. After termination of the reaction, 5 μl of the reaction mixture was electrophoresed on 1% agarose gel to confirm amplified fragments.

[0194] As a result, there can be confirmed that a DNA fragment of 1 kb was most efficiently amplified, in the case of the system where the rRFC-M complex and the F7 protein were added, as a result of comparison of the system where the rRFC-M complex and the F7 protein were added with the system where the rRFC-M complex alone was added, the system where the F7 protein alone was added, or the system where LA Taq DNA polymerase alone was added.

INDUSTRIAL APPLICABILITY

[0195] According to the present invention, there can be provided a DNA polymerase-associated factor capable of enhancing DNA synthesizing-activity of a DNA polymerase. The factor has an action on various DNA polymerases, and also can be utilized in various processes in which a DNA polymerase is used, so that the factor is useful as a reagent for studies in genetic engineering. Further, it is now possible to produce the enzyme by genetic engineering techniques using a gene encoding the DNA polymerase-associated factor of the present invention.

SEQUENCE LISTING

SEQ ID NO: 1
SEQUENCE LENGTH: 249
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Pro Phe Glu Ile Val Phe Glu Gly Ala Lys Glu Phe Ala Gln
                5                   10                  15
Leu Ile Asp Thr Ala Ser Lys Leu Ile Asp Glu Ala Ala Phe Lys
                20                  25                  30
Val Thr Glu Asp Gly Ile Ser Met Arg Ala Met Asp Pro Ser Arg
                35                  40                  45
Val Val Leu Ile Asp Leu Asn Leu Pro Ser Ser Ile Phe Ser Lys
                50                  55                  60
Tyr Glu Val Val Glu Pro Glu Thr Ile Gly Val Asn Met Asp His
                65                  70                  75
Leu Lys Lys Ile Leu Lys Arg Gly Lys Ala Lys Asp Thr Leu Ile
                80                  85                  90
Leu Lys Lys Gly Glu Glu Asn Phe Leu Glu Ile Thr Ile Gln Gly
                95                  100                 105
Thr Ala Thr Arg Thr Phe Arg Val Pro Leu Ile Asp Val Glu Glu
                110                 115                 120
Met Glu Val Asp Leu Pro Glu Leu Pro Phe Thr Ala Lys Val Val
                125                 130                 135
Val Leu Gly Glu Val Leu Lys Asp Ala Val Lys Asp Ala Ser Leu
                140                 145                 150
Val Ser Asp Ser Ile Lys Phe Ile Ala Arg Glu Asn Glu Phe Ile
                155                 160                 165
Met Lys Ala Glu Gly Glu Thr Gln Glu Val Glu Ile Lys Leu Thr
                170                 175                 180
Leu Glu Asp Glu Gly Leu Leu Asp Ile Glu Val Gln Glu Glu Thr
```

```
                185                     190                     195
Lys Ser Ala Tyr Gly Val Ser Tyr Leu Ser Asp Met Val Lys Gly
                200                     205                     210
Leu Gly Lys Ala Asp Glu Val Thr Ile Lys Phe Gly Asn Glu Met
                215                     220                     225
Pro Met Gln Met Glu Tyr Tyr Ile Arg Asp Glu Gly Arg Leu Thr
                230                     235                     240
Phe Leu Leu Ala Pro Arg Val Glu Glu
                245
```

SEQ ID NO: 2
SEQUENCE LENGTH: 750
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:

```
ATGCCATTTG AAATCGTATT TGAAGGTGCA AAAGAGTTTG CCCAACTTAT AGACACCGCA    60
AGTAAGTTAA TAGATGAGGC CGCGTTTAAA GTTACAGAAG ATGGGATAAG CATGAGGGCC   120
ATGGATCCAA GTAGAGTTGT CCTGATTGAC CTAAATCTCC CGTCAAGCAT ATTTAGCAAA   180
TATGAAGTTG TTGAACCAGA AACAATTGGA GTTAACATGG ACCACCTAAA GAAGATCCTA   240
AAGAGAGGTA AAGCAAAGGA CACCTTAATA CTCAAGAAAG GAGAGGAAAA CTTCTTAGAG   300
ATAACAATTC AAGGAACTGC AACAAGAACA TTTAGAGTTC CCCTAATAGA TGTAGAAGAG   360
ATGGAAGTTG ACCTCCCAGA ACTTCCATTC ACTGCAAAGG TTGTAGTTCT TGGAGAAGTC   420
CTAAAAGATG CTGTTAAAGA TGCCTCTCTA GTGAGTGACA GCATAAAATT TATTGCCAGG   480
GAAAATGAAT TTATAATGAA GGCAGAGGGA GAAACCCAGG AAGTTGAGAT AAAGCTAACT   540
CTTGAAGATG AGGGATTATT GGACATCGAG GTTCAAGAGG AGACAAAGAG CGCATATGGA   600
GTCAGCTATC TCTCCGACAT GGTTAAAGGA CTTGGAAAGG CCGATGAAGT TACAATAAAG   660
TTTGGAAATG AAATGCCCAT GCAAATGGAG TATTACATTA GAGATGAAGG AAGACTTACA   720
TTCCTACTGG CTCCAAGAGT TGAAGAGTGA                                    750
```

SEQ ID NO: 3
SEQUENCE LENGTH: 327
SEQUENCE TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Ser Glu Glu Ile Arg Glu Val Lys Val Leu Glu Lys Pro Trp
                  5                  10                  15
Val Glu Lys Tyr Arg Pro Gln Arg Leu Asp Asp Ile Val Gly Gln
                 20                  25                  30
Glu His Ile Val Lys Arg Leu Lys His Tyr Val Lys Thr Gly Ser
                 35                  40                  45
Met Pro His Leu Leu Phe Ala Gly Pro Pro Gly Val Gly Lys Thr
                 50                  55                  60
Thr Ala Ala Leu Ala Leu Ala Arg Glu Leu Phe Gly Glu Asn Trp
                 65                  70                  75
Arg His Asn Phe Leu Glu Leu Asn Ala Ser Asp Glu Arg Gly Ile
                 80                  85                  90
Asn Val Ile Arg Glu Lys Val Lys Glu Phe Ala Arg Thr Lys Pro
                 95                 100                 105
Ile Gly Gly Ala Ser Phe Lys Ile Ile Phe Leu Asp Glu Ala Asp
                110                 115                 120
Ala Leu Thr Gln Asp Ala Gln Gln Ala Leu Arg Arg Thr Met Glu
                125                 130                 135
Met Phe Ser Ser Asn Val Arg Phe Ile Leu Ser Cys Asn Tyr Ser
                140                 145                 150
Ser Lys Ile Ile Glu Pro Ile Gln Ser Arg Cys Ala Ile Phe Arg
                155                 160                 165
Phe Arg Pro Leu Arg Asp Glu Asp Ile Ala Lys Arg Leu Arg Tyr
                170                 175                 180
Ile Ala Glu Asn Glu Gly Leu Glu Leu Thr Glu Glu Gly Leu Gln
                185                 190                 195
Ala Ile Leu Tyr Ile Ala Glu Gly Asp Met Arg Arg Ala Ile Asn
                200                 205                 210
Ile Leu Gln Ala Ala Ala Ala Leu Asp Lys Lys Ile Thr Asp Glu
                215                 220                 225
```

Asn Val Phe Met Val Ala Ser Arg Ala Arg Pro Glu Asp Ile Arg
230 235 240

Glu Met Met Leu Leu Ala Leu Lys Gly Asn Phe Leu Lys Ala Arg
245 250 255

Glu Lys Leu Arg Glu Ile Leu Leu Lys Gln Gly Leu Ser Gly Glu
260 265 270

Asp Val Leu Val Gln Met His Lys Glu Val Phe Asn Leu Pro Ile
275 280 285

Glu Glu Pro Lys Lys Val Leu Leu Ala Asp Lys Ile Gly Glu Tyr
290 295 300

Asn Phe Arg Leu Val Glu Gly Ala Asn Glu Ile Ile Gln Leu Glu
305 310 315

Ala Leu Leu Ala Gln Phe Thr Leu Ile Gly Lys Lys
320 325

SEQ ID NO: 4
SEQUENCE LENGTH: 984
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:

ATGAGCGAAG AGATTAGAGA AGTTAAGGTT CTAGAAAAAC CCTGGGTTGA GAAGTATAGA 60
CCTCAAAGAC TTGACGACAT TGTAGGACAA GAGCACATAG TGAAAAGGCT CAAGCACTAC 120
GTCAAAACTG GATCAATGCC CCACCTACTC TTCGCAGGCC CCCCTGGTGT CGGAAAGACT 180
ACAGCGGCTT TGGCCCTTGC AAGAGAGCTT TTCGGCGAAA ACTGGAGGCA TAACTTCCTC 240
GAGTTGAATG CTTCAGATGA AAGAGGTATA AACGTAATTA GAGAGAAAGT TAAGGAGTTT 300
GCGAGAACAA AGCCTATAGG AGGAGCAAGC TTCAAGATAA TTTTCCTTGA TGAGGCCGAC 360
GCTTTAACTC AAGATGCCCA ACAAGCCTTA AGAAGAACCA TGGAAATGTT CTCGAGTAAC 420
GTTCGCTTTA TCTTGAGCTG TAACTACTCC TCCAAGATAA TTGAACCCAT ACAGTCTAGA 480
TGTGCAATAT TCCGCTTCAG ACCTCTCCGC GATGAGGATA TAGCGAAGAG ACTAAGGTAC 540
ATTGCCGAAA ATGAGGGCTT AGAGCTAACT GAAGAAGGTC TCCAAGCAAT ACTTTACATA 600
GCAGAAGGAG ATATGAGAAG AGCAATAAAC ATTCTGCAAG CTGCAGCAGC TCTAGACAAG 660
AAGATCACCG ACGAAAACGT ATTCATGGTA GCGAGTAGAG CTAGACCTGA AGATATAAGA 720

GAGATGATGC TTCTTGCTCT CAAAGGCAAC TTCTTGAAGG CCAGAGAAAA GCTTAGGGAG 780
ATACTTCTCA AGCAAGGACT TAGTGGAGAA GATGTACTAG TTCAGATGCA CAAAGAAGTC 840
TTCAACCTGC CAATAGAGGA GCCAAAGAAG GTTCTGCTTG CTGATAAGAT AGGAGAGTAT 900
AACTTCAGAC TCGTTGAAGG GGCTAATGAA ATAATTCAGC TTGAAGCACT CTTAGCACAG 960
TTCACCCTAA TTGGGAAGAA GTGA 984

SEQ ID NO: 5
SEQUENCE LENGTH: 613
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

Met Asp Glu Phe Val Lys Ser Leu Leu Lys Ala Asn Tyr Leu Ile
5 10 15
Thr Pro Ser Ala Tyr Tyr Leu Leu Arg Glu Tyr Tyr Glu Lys Gly
20 25 30
Glu Phe Ser Ile Val Glu Leu Val Lys Phe Ala Arg Ser Arg Glu
35 40 45
Ser Tyr Ile Ile Thr Asp Ala Leu Ala Thr Glu Phe Leu Lys Val
50 55 60
Lys Gly Leu Glu Pro Ile Leu Pro Val Glu Thr Lys Gly Gly Phe
65 70 75
Val Ser Thr Gly Glu Ser Gln Lys Glu Gln Ser Tyr Glu Glu Ser
80 85 90
Phe Gly Thr Lys Glu Glu Ile Ser Gln Glu Ile Lys Glu Gly Glu
95 100 105
Ser Phe Ile Ser Thr Gly Ser Glu Pro Leu Glu Glu Glu Leu Asn
110 115 120
Ser Ile Gly Ile Glu Glu Ile Gly Ala Asn Glu Glu Leu Val Ser
125 130 135
Asn Gly Asn Asp Asn Gly Gly Glu Ala Ile Val Phe Asp Lys Tyr
140 145 150

Gly Tyr Pro Met Val Tyr Ala Pro Glu Glu Ile Glu Val Glu Glu
155 160 165

Lys Glu Tyr Ser Lys Tyr Glu Asp Leu Thr Ile Pro Met Asn Pro
170 175 180

Asp Phe Asn Tyr Val Glu Ile Lys Glu Asp Tyr Asp Val Val Phe
185 190 195

Asp Val Arg Asn Val Lys Leu Lys Pro Pro Lys Val Lys Asn Gly
200 205 210

Asn Gly Lys Glu Gly Glu Ile Ile Val Glu Ala Tyr Ala Ser Leu
215 220 225

Phe Arg Ser Arg Leu Lys Lys Leu Arg Lys Ile Leu Arg Glu Asn
230 235 240

Pro Glu Leu Asp Asn Val Val Asp Ile Gly Lys Leu Lys Tyr Val
245 250 255

Lys Glu Asp Glu Thr Val Thr Ile Ile Gly Leu Val Asn Ser Lys
260 265 270

Arg Glu Val Asn Lys Gly Leu Ile Phe Glu Ile Glu Asp Leu Thr
275 280 285

Gly Lys Val Lys Val Phe Leu Pro Lys Asp Ser Glu Asp Tyr Arg
290 295 300

Glu Ala Phe Lys Val Leu Pro Asp Ala Val Val Ala Phe Lys Gly
305 310 315

Val Tyr Ser Lys Arg Gly Ile Leu Tyr Ala Asn Lys Phe Tyr Leu
320 325 330

Pro Asp Val Pro Leu Tyr Arg Arg Gln Lys Pro Pro Leu Glu Glu
335 340 345

Lys Val Tyr Ala Ile Leu Ile Ser Asp Ile His Val Gly Ser Lys
350 355 360

Glu Phe Cys Glu Asn Ala Phe Ile Lys Phe Leu Glu Trp Leu Asn
365 370 375

Gly Asn Val Glu Thr Lys Glu Glu Glu Glu Ile Val Ser Arg Val
380 385 390

Lys Tyr Leu Ile Ile Ala Gly Asp Val Val Asp Gly Val Gly Val
395 400 405

Tyr Pro Gly Gln Tyr Ala Asp Leu Thr Ile Pro Asp Ile Phe Asp
410 415 420

Gln Tyr Glu Ala Leu Ala Asn Leu Leu Ser His Val Pro Lys His
425 430 435

Ile Thr Met Phe Ile Ala Pro Gly Asn His Asp Ala Ala Arg Gln
440 445 450

Ala Ile Pro Gln Pro Glu Phe Tyr Lys Glu Tyr Ala Lys Pro Ile
455 460 465

Tyr Lys Leu Lys Asn Ala Val Ile Ile Ser Asn Pro Ala Val Ile
470 475 480

Arg Leu His Gly Arg Asp Phe Leu Ile Ala His Gly Arg Gly Ile
485 490 495

Glu Asp Val Val Gly Ser Val Pro Gly Leu Thr His His Lys Pro
500 505 510

Gly Leu Pro Met Val Glu Leu Leu Lys Met Arg His Val Ala Pro
515 520 525

Met Phe Gly Gly Lys Val Pro Ile Ala Pro Asp Pro Glu Asp Leu
530 535 540

Leu Val Ile Glu Glu Val Pro Asp Val Val His Met Gly His Val
545 550 555

His Val Tyr Asp Ala Val Val Tyr Arg Gly Val Gln Leu Val Asn
560 565 570

Ser Ala Thr Trp Gln Ala Gln Thr Glu Phe Gln Lys Met Val Asn
575 580 585

Ile Val Pro Thr Pro Ala Lys Val Pro Val Val Asp Ile Asp Thr
590 595 600

Ala Lys Val Val Lys Val Leu Asp Phe Ser Gly Trp Cys
605 610

SEQ ID NO: 6
SEQUENCE LENGTH: 1263
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULAR TYPE: peptide

SEQUENCE DESCRIPTION:

```
Met Glu Leu Pro Lys Glu Ile Glu Glu Tyr Phe Glu Met Leu Gln
                  5                  10                  15
Arg Glu Ile Asp Lys Ala Tyr Glu Ile Ala Lys Lys Ala Arg Ser
                 20                  25                  30
Gln Gly Lys Asp Pro Ser Thr Asp Val Glu Ile Pro Gln Ala Thr
                 35                  40                  45
Asp Met Ala Gly Arg Val Glu Ser Leu Val Gly Pro Pro Gly Val
                 50                  55                  60
Ala Gln Arg Ile Arg Glu Leu Leu Lys Glu Tyr Asp Lys Glu Ile
                 65                  70                  75
Val Ala Leu Lys Ile Val Asp Glu Ile Ile Glu Gly Lys Phe Gly
                 80                  85                  90
Asp Phe Gly Ser Lys Glu Lys Tyr Ala Glu Gln Ala Val Arg Thr
                 95                 100                 105
Ala Leu Ala Ile Leu Thr Glu Gly Ile Val Ser Ala Pro Leu Glu
                110                 115                 120
Gly Ile Ala Asp Val Lys Ile Lys Arg Asn Thr Trp Ala Asp Asn
                125                 130                 135
Ser Glu Tyr Leu Ala Leu Tyr Tyr Ala Gly Pro Ile Arg Ser Ser
                140                 145                 150
Gly Gly Thr Ala Gln Ala Leu Ser Val Leu Val Gly Asp Tyr Val
                155                 160                 165
Arg Arg Lys Leu Gly Leu Asp Arg Phe Lys Pro Ser Gly Lys His
                170                 175                 180
Ile Glu Arg Met Val Glu Glu Val Asp Leu Tyr His Arg Ala Val
                185                 190                 195
Ser Arg Leu Gln Tyr His Pro Ser Pro Asp Glu Val Arg Leu Ala
                200                 205                 210
Met Arg Asn Ile Pro Ile Glu Ile Thr Gly Glu Ala Thr Asp Asp
                215                 220                 225
Val Glu Val Ser His Arg Asp Val Glu Gly Val Glu Thr Asn Gln
                230                 235                 240
```

Leu Arg Gly Gly Ala Ile Leu Val Leu Ala Glu Gly Val Leu Gln
245 250 255

Lys Ala Lys Lys Leu Val Lys Tyr Ile Asp Lys Met Gly Ile Asp
260 265 270

Gly Trp Glu Trp Leu Lys Glu Phe Val Glu Ala Lys Glu Lys Gly
275 280 285

Glu Glu Ile Glu Glu Ser Glu Ser Lys Ala Glu Glu Ser Lys Val
290 295 300

Glu Thr Arg Val Glu Val Glu Lys Gly Phe Tyr Tyr Lys Leu Tyr
305 310 315

Glu Lys Phe Arg Ala Glu Ile Ala Pro Ser Glu Lys Tyr Ala Lys
320 325 330

Glu Ile Ile Gly Gly Arg Pro Leu Phe Ala Gly Pro Ser Glu Asn
335 340 345

Gly Gly Phe Arg Leu Arg Tyr Gly Arg Ser Arg Val Ser Gly Phe
350 355 360

Ala Thr Trp Ser Ile Asn Pro Ala Thr Met Val Leu Val Asp Glu
365 370 375

Phe Leu Ala Ile Gly Thr Gln Met Lys Thr Glu Arg Pro Gly Lys
380 385 390

Gly Ala Val Val Thr Pro Ala Thr Thr Ala Glu Gly Pro Ile Val
395 400 405

Lys Leu Lys Asp Gly Ser Val Val Arg Val Asp Asp Tyr Asn Leu
410 415 420

Ala Leu Lys Ile Arg Asp Glu Val Glu Glu Ile Leu Tyr Leu Gly
425 430 435

Asp Ala Ile Ile Ala Phe Gly Asp Phe Val Glu Asn Asn Gln Thr
440 445 450

Leu Leu Pro Ala Asn Tyr Val Glu Glu Trp Trp Ile Gln Glu Phe
455 460 465

Val Lys Ala Val Asn Glu Ala Tyr Glu Val Glu Leu Arg Pro Phe
470 475 480

Glu Glu Asn Pro Arg Glu Ser Val Glu Glu Ala Ala Glu Tyr Leu
485 490 495

Glu Val Asp Pro Glu Phe Leu Ala Lys Met Leu Tyr Asp Pro Leu
500 505 510

Arg Val Lys Pro Pro Val Glu Leu Ala Ile His Phe Ser Glu Ile
515 520 525

Leu Glu Ile Pro Leu His Pro Tyr Tyr Thr Leu Tyr Trp Asn Thr
530 535 540

Val Asn Pro Lys Asp Val Glu Arg Leu Trp Gly Val Leu Lys Asp
545 550 555

Lys Ala Thr Ile Glu Trp Gly Thr Phe Arg Gly Ile Lys Phe Ala
560 565 570

Lys Lys Ile Glu Ile Ser Leu Asp Asp Leu Gly Ser Leu Lys Arg
575 580 585

Thr Leu Glu Leu Leu Gly Leu Pro His Thr Val Arg Glu Gly Ile
590 595 600

Val Val Val Asp Tyr Pro Trp Ser Ala Ala Leu Leu Thr Pro Leu
605 610 615

Gly Asn Leu Glu Trp Glu Phe Lys Ala Lys Pro Phe Tyr Thr Val
620 625 630

Ile Asp Ile Ile Asn Glu Asn Asn Gln Ile Lys Leu Arg Asp Arg
635 640 645

Gly Ile Ser Trp Ile Gly Ala Arg Met Gly Arg Pro Glu Lys Ala
650 655 660

Lys Glu Arg Lys Met Lys Pro Pro Val Gln Val Leu Phe Pro Ile
665 670 675

Gly Leu Ala Gly Gly Ser Ser Arg Asp Ile Lys Lys Ala Ala Glu
680 685 690

Glu Gly Lys Ile Ala Glu Val Glu Ile Ala Phe Phe Lys Cys Pro
695 700 705

Lys Cys Gly His Val Gly Pro Glu Thr Leu Cys Pro Glu Cys Gly
710 715 720

Ile Arg Lys Glu Leu Ile Trp Thr Cys Pro Lys Cys Gly Ala Glu
725 730 735

Tyr Thr Asn Ser Gln Ala Glu Gly Tyr Ser Tyr Ser Cys Pro Lys
740 745 750

Cys Asn Val Lys Leu Lys Pro Phe Thr Lys Arg Lys Ile Lys Pro
755 760 765

Ser Glu Leu Leu Asn Arg Ala Met Glu Asn Val Lys Val Tyr Gly
770 775 780

Val Asp Lys Leu Lys Gly Val Met Gly Met Thr Ser Gly Trp Lys
785 790 795

Ile Ala Glu Pro Leu Glu Lys Gly Leu Leu Arg Ala Lys Asn Glu
800 805 810

Val Tyr Val Phe Lys Asp Gly Thr Ile Arg Phe Asp Ala Thr Asp
815 820 825

Ala Pro Ile Thr His Phe Arg Pro Arg Glu Ile Gly Val Ser Val
830 835 840

Glu Lys Leu Arg Glu Leu Gly Tyr Thr His Asp Phe Glu Gly Lys
845 850 855

Pro Leu Val Ser Glu Asp Gln Ile Val Glu Leu Lys Pro Gln Asp
860 865 870

Val Ile Leu Ser Lys Glu Ala Gly Lys Tyr Leu Leu Arg Val Ala
875 880 885

Arg Phe Val Asp Asp Leu Leu Glu Lys Phe Tyr Gly Leu Pro Arg
890 895 900

Phe Tyr Asn Ala Glu Lys Met Glu Asp Leu Ile Gly His Leu Val
905 910 915

Ile Gly Leu Ala Pro His Thr Ser Ala Gly Ile Val Gly Arg Ile
920 925 930

Ile Gly Phe Val Asp Ala Leu Val Gly Tyr Ala His Pro Tyr Phe
935 940 945

His Ala Ala Lys Arg Arg Asn Cys Asp Gly Asp Glu Asp Ser Val
950 955 960

Met Leu Leu Leu Asp Ala Leu Leu Asn Phe Ser Arg Tyr Tyr Leu
965 970 975

Pro Glu Lys Arg Gly Gly Lys Met Asp Ala Pro Leu Val Ile Thr
980 985 990

Thr Arg Leu Asp Pro Arg Glu Val Asp Ser Glu Val His Asn Met
995 1000 1005

Asp Val Val Arg Tyr Tyr Pro Leu Glu Phe Tyr Glu Ala Thr Tyr
1010 1015 1020

Glu Leu Lys Ser Pro Lys Glu Leu Val Arg Val Ile Glu Gly Val
1025 1030 1035

Glu Asp Arg Leu Gly Lys Pro Glu Met Tyr Tyr Gly Ile Lys Phe
1040 1045 1050

Thr His Asp Thr Asp Asp Ile Ala Leu Gly Pro Lys Met Ser Leu
1055 1060 1065

Tyr Lys Gln Leu Gly Asp Met Glu Glu Lys Val Lys Arg Gln Leu
1070 1075 1080

Thr Leu Ala Glu Arg Ile Arg Ala Val Asp Gln His Tyr Val Ala
1085 1090 1095

Glu Thr Ile Leu Asn Ser His Leu Ile Pro Asp Leu Arg Gly Asn
1100 1105 1110

Leu Arg Ser Phe Thr Arg Gln Glu Phe Arg Cys Val Lys Cys Asn
1115 1120 1125

Thr Lys Tyr Arg Arg Pro Pro Leu Asp Gly Lys Cys Pro Val Cys
1130 1135 1140

Gly Gly Lys Ile Val Leu Thr Val Ser Lys Gly Ala Ile Glu Lys
1145 1150 1155

Tyr Leu Gly Thr Ala Lys Met Leu Val Ala Asn Tyr Asn Val Lys
1160 1165 1170

Pro Tyr Thr Arg Gln Arg Ile Cys Leu Thr Glu Lys Asp Ile Asp
1175 1180 1185

Ser Leu Phe Glu Tyr Leu Phe Pro Glu Ala Gln Leu Thr Leu Ile
1190 1195 1200

Val Asp Pro Asn Asp Ile Cys Met Lys Met Ile Lys Glu Arg Thr
1205 1210 1215

Gly Glu Thr Val Gln Gly Gly Leu Leu Glu Asn Phe Asn Ser Ser
1220 1225 1230

Gly Asn Asn Gly Lys Lys Ile Glu Lys Lys Glu Lys Lys Ala Lys
1235 1240 1245

Glu Lys Pro Lys Lys Lys Lys Val Ile Ser Leu Asp Asp Phe Phe
1250 1255 1260

Ser Lys Arg

SEQ ID NO: 7
SEQUENCE LENGTH: 20
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Asp Lys Glu Gly Phe Leu Asn Lys Val Arg Glu Ala Val Asp
                  5                  10                  15
Val Val Lys Leu His
                 20
```

SEQ ID NO: 8
SEQUENCE LENGTH: 20
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Phe Thr Gly Lys Val Leu Ile Pro Val Lys Val Leu Lys Lys
                  5                  10                  15
Phe Glu Asn Trp Asn
                 20
```

SEQ ID NO: 9
SEQUENCE LENGTH: 20
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Ile Gly Ser Ile Phe Tyr Ser Lys Lys Phe Asn Leu His Arg
```

5 10 15

Pro Ser Glu Tyr His

20

SEQ ID NO: 10

SEQUENCE LENGTH: 20

SEQUENCE TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULAR TYPE: peptide

SEQUENCE DESCRIPTION:

Met Lys Asp Tyr Arg Pro Leu Leu Gly Ala Ile Lys Val Lys Gly

5 10 15

Asp Asn Val Phe Ser

20

SEQ ID NO: 11

SEQUENCE LENGTH: 18

SEQUENCE TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULAR TYPE: peptide

SEQUENCE DESCRIPTION:

Met Asp Ile Glu Val Leu Arg Arg Leu Leu Glu Arg Glu Leu Ser

5 10 15

Ser Glu His

SEQ ID NO: 12

SEQUENCE LENGTH: 17

SEQUENCE TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULAR TYPE: peptide

SEQUENCE DESCRIPTION:

Pro Phe Glu Ile Val Phe Glu Gly Ala Lys Glu Phe Ala Gln Leu
5 10 15
Ile Asp

SEQ ID NO: 13
SEQUENCE LENGTH: 17
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
ATGGATAARG ARGGNTT 17

SEQ ID NO: 14
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
AATAAAGTWA GRGARGCNGT 20

SEQ ID NO: 15
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CTCTGCGGCA ATTCTTGCAA 20

SEQ ID NO: 16
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CTTGCAAAGA AGTATGTAAC 20

SEQ ID NO: 17
SEQUENCE LENGTH: 2009
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:

```
AAGCTTCCAA AGAACTGGCG TTACGACCCA GAGACTGCAA AGTTGCTCGT CCGCTGATCC   60
TTCCCTATAT TTTCATTTGG TGTTTTTCAT GGATAAGGAG GGTTTTTTGA ACAAGGTTAG  120
GGAGGCTGTG GATGTAGTAA AGCTCCACAT CGAGTTAGGT CATACTATAA GGATAATCTC  180
TCATAGGGAT GCGGATGGAA TAACCTCTGC GGCAATTCTT GCAAAGGCTT TGGGAAGAGA  240
AGGAGCGAGC TTTCACATTT CGATTGTTAA ACAGGTAAGT GAAGATCTTT TAAGAGAATT  300
AAAGGATGAA GATTACAAAA TCTTCATTTT TTCCGACCTG GGTAGTGGTT CTTTAAGTTT  360
GATAAAAGAG TATCTTAAGG AAAAAACTGT TATAATCCTT GATCACCATC CTCCGGAAAA  420
TGTGAAGTTG GAAGAAAAGC ATATACTTGT TAATCCAGTT CAATTTGGCG CAAATAGCGT  480
TAGGGATCTG AGTGGATCTG GGGTTACATA CTTCTTTGCA AGGGAGCTAA ATGAAAAGAA  540
TAGGGACCTT GCTTACATTG CAATAGTGGG AGCAGTTGGG GATATGCAAG AGAACGATGG  600
AGTTTTCCAT GGGATGAACC TTGATATTAT TGAAGATGGG AAATCTCTGG GAATTCTTGA  660
GGTTAAAAAA GAATTGCGCC TGTTTGGTAG GGAAACTAGA CCTCTCTATC AAATGCTCGC  720
ATATGCCACA AATCCGGAAA TTCCTGAAGT TACTGGAGAC GAGAGGAAGG CCATAGAGTG  780
GTTAAAGAAC AAGGGCTTCA ATCCCGAGAA AAAATATTGG GAATTAAGTG AGGAGGAAAA  840
GAAAAAGTTA CATGATTTCC TAATCATTCA CATGATCAAG CATGGAGCTG GAAAAGAGGA  900
TATAGATAGG CTAATAGGAG ACGTTGTTAT TAGTCCCTTA TATCCTGAAG GGGATCCCAG  960
GCACGAGGCT AGAGAATTTG CTACCCTATT AAACGCTACA GGCAGGTTAA ACTTGGGCAA 1020
CTTAGGAGTG GCTGTATGTT TGGGAGATGA GGAGGCTTTC AGAAAGGCCC TAAAGATGGT 1080
TGAAGACTAC AAGAGGGAGC AAATTGAAGC AAGAAAGTGG CTACTTCAAA ATTGGAACAG 1140
TGAAGTTTGG GAGGGGGATC ATGTTTACGT CTTATATGTG GGAAAGAGTA TTAGAGATAC 1200
TCTCGTTGGA ATAGCAGCTA GCATGGCCAT CAATGCTGGA CTGGCAGATC CTGAAAAGCC 1260
```

GGTTATAGTG TTTGCAGATA CTGATGAAGA TCCAAACCTT CTCAAAGGTT CAGCTAGAAC 1320
AACTGAAAGG GCTTTAGCTA AGGGTTACAA TTTGGGAGAA GCTCTTAGGA AAGCGGCTGA 1380
GCTAGTGAAT GGGGAAGGGG GAGGACACGC GATAGCTGCA GGTATAAGAA TTCCCAGGGC 1440
CAGGTTGGCG GAGTTTAGAA AATTAATAGA TAAAATCCTT GGAGAACAGG TGAGCAAAGG 1500
TGGAGATAAA AGCGAAAGCT GAAATATTGT GGGAGTACAG CGATGAGAAG GTTGCTGAGG 1560
CTATTGCGAA GTCTGTTGAT GTTGATAATA TTTCTCTCCC TCCAAACCTC AAGAAAAGTT 1620
TAAATCTTAT GACGTTTTCC GATGGAGCGA AGGTAATAAC AAAGGTTAAA TATCATGGAG 1680
AAATTGAGAC TCTCATAGTT GCTCTCGATG ATTTGATATT CGCTGTAAAA GTTGCTGAGG 1740
AGGTGTTATG ATGGTGNGAA AAGGGNAACA ACAACANGGG ATAAGGGAAG NTGAAGCAAT 1800
GGTATATTAT TTATGCTCCN GANTTCTTGG GCGGGGTAGA GGTAGGATTA ACGCCAGCAG 1860
ACGATCCAGA GAAAGTACTC AACAGAGTCG TTGAAGTTAC TCTGAAGGAT GTTACAGGAG 1920
ACTTTACAAA GAGTCACGTG AAGCTCTATT TCCAAGTATA TGATGTCAAG GGACAGAATG 1980
CCTACACAAA GTTCAAGGGA ATGAAGCTT 2009

SEQ ID NO: 18
SEQUENCE LENGTH: 1434
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:

ATGGATAAGG AGGGTTTTTT GAACAAGGTT AGGGAGGCTG TGGATGTAGT AAAGCTCCAC 60
ATCGAGTTAG GTCATACTAT AAGGATAATC TCTCATAGGG ATGCGGATGG AATAACCTCT 120
GCGGCAATTC TTGCAAAGGC TTTGGGAAGA GAAGGAGCGA GCTTTCACAT TTCGATTGTT 180
AAACAGGTAA GTGAAGATCT TTTAAGAGAA TTAAAGGATG AAGATTACAA AATCTTCATT 240
TTTTCCGACC TGGGTAGTGG TTCTTTAAGT TTGATAAAAG AGTATCTTAA GGAAAAAACT 300
GTTATAATCC TTGATCACCA TCCTCCGGAA AATGTGAAGT TGGAAGAAAA GCATATACTT 360
GTTAATCCAG TTCAATTTGG CGCAAATAGC GTTAGGGATC TGAGTGGATC TGGGGTTACA 420
TACTTCTTTG CAAGGGAGCT AAATGAAAAG AATAGGGACC TTGCTTACAT TGCAATAGTG 480
GGAGCAGTTG GGGATATGCA AGAGAACGAT GGAGTTTTCC ATGGGATGAA CCTTGATATT 540
ATTGAAGATG GGAAATCTCT GGGAATTCTT GAGGTTAAAA AAGAATTGCG CCTGTTTGGT 600
AGGGAAACTA GACCTCTCTA TCAAATGCTC GCATATGCCA CAAATCCGGA AATTCCTGAA 660
GTTACTGGAG ACGAGAGGAA GGCCATAGAG TGGTTAAAGA ACAAGGGCTT CAATCCCGAG 720
AAAAAATATT GGGAATTAAG TGAGGAGGAA AAGAAAAAGT TACATGATTT CCTAATCATT 780

CACATGATCA AGCATGGAGC TGGAAAAGAG GATATAGATA GGCTAATAGG AGACGTTGTT 840
ATTAGTCCCT TATATCCTGA AGGGGATCCC AGGCACGAGG CTAGAGAATT TGCTACCCTA 900
TTAAACGCTA CAGGCAGGTT AAACTTGGGC AACTTAGGAG TGGCTGTATG TTTGGGAGAT 960
GAGGAGGCTT TCAGAAAGGC CCTAAAGATG GTTGAAGACT ACAAGAGGGA GCAAATTGAA 1020
GCAAGAAAGT GGCTACTTCA AAATTGGAAC AGTGAAGTTT GGGAGGGGGA TCATGTTTAC 1080
GTCTTATATG TGGGAAAGAG TATTAGAGAT ACTCTCGTTG GAATAGCAGC TAGCATGGCC 1140
ATCAATGCTG GACTGGCAGA TCCTGAAAAG CCGGTTATAG TGTTTGCAGA TACTGATGAA 1200
GATCCAAACC TTCTCAAAGG TTCAGCTAGA ACAACTGAAA GGGCTTTAGC TAAGGGTTAC 1260
AATTTGGGAG AAGCTCTTAG GAAAGCGGCT GAGCTAGTGA ATGGGGAAGG GGGAGGACAC 1320
GCGATAGCTG CAGGTATAAG AATTCCCAGG GCCAGGTTGG CGGAGTTTAG AAAATTAATA 1380
GATAAAATCC TTGGAGAACA GGTGAGCAAA GGTGGAGATA AAAGCGAAAG CTGA 1434

SEQ ID NO: 19
SEQUENCE LENGTH: 477
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

Met Asp Lys Glu Gly Phe Leu Asn Lys Val Arg Glu Ala Val Asp
5 10 15
Val Val Lys Leu His Ile Glu Leu Gly His Thr Ile Arg Ile Ile
20 25 30
Ser His Arg Asp Ala Asp Gly Ile Thr Ser Ala Ala Ile Leu Ala
35 40 45
Lys Ala Leu Gly Arg Glu Gly Ala Ser Phe His Ile Ser Ile Val
50 55 60
Lys Gln Val Ser Glu Asp Leu Leu Arg Glu Leu Lys Asp Glu Asp
65 70 75
Tyr Lys Ile Phe Ile Phe Ser Asp Leu Gly Ser Gly Ser Leu Ser
80 85 90
Leu Ile Lys Glu Tyr Leu Lys Glu Lys Thr Val Ile Ile Leu Asp
95 100 105
His His Pro Pro Glu Asn Val Lys Leu Glu Glu Lys His Ile Leu

110 115 120
Val Asn Pro Val Gln Phe Gly Ala Asn Ser Val Arg Asp Leu Ser
125 130 135
Gly Ser Gly Val Thr Tyr Phe Phe Ala Arg Glu Leu Asn Glu Lys
140 145 150
Asn Arg Asp Leu Ala Tyr Ile Ala Ile Val Gly Ala Val Gly Asp
155 160 165
Met Gln Glu Asn Asp Gly Val Phe His Gly Met Asn Leu Asp Ile
170 175 180
Ile Glu Asp Gly Lys Ser Leu Gly Ile Leu Glu Val Lys Lys Glu
185 190 195
Leu Arg Leu Phe Gly Arg Glu Thr Arg Pro Leu Tyr Gln Met Leu
200 205 210
Ala Tyr Ala Thr Asn Pro Glu Ile Pro Glu Val Thr Gly Asp Glu
215 220 225
Arg Lys Ala Ile Glu Trp Leu Lys Asn Lys Gly Phe Asn Pro Glu
230 235 240
Lys Lys Tyr Trp Glu Leu Ser Glu Glu Glu Lys Lys Lys Leu His
245 250 255
Asp Phe Leu Ile Ile His Met Ile Lys His Gly Ala Gly Lys Glu
260 265 270
Asp Ile Asp Arg Leu Ile Gly Asp Val Val Ile Ser Pro Leu Tyr
275 280 285
Pro Glu Gly Asp Pro Arg His Glu Ala Arg Glu Phe Ala Thr Leu
290 295 300
Leu Asn Ala Thr Gly Arg Leu Asn Leu Gly Asn Leu Gly Val Ala
305 310 315
Val Cys Leu Gly Asp Glu Glu Ala Phe Arg Lys Ala Leu Lys Met
320 325 330
Val Glu Asp Tyr Lys Arg Glu Gln Ile Glu Ala Arg Lys Trp Leu
335 340 345
Leu Gln Asn Trp Asn Ser Glu Val Trp Glu Gly Asp His Val Tyr
350 355 360
Val Leu Tyr Val Gly Lys Ser Ile Arg Asp Thr Leu Val Gly Ile

```
                    365                     370                     375
Ala Ala Ser Met Ala Ile Asn Ala Gly Leu Ala Asp Pro Glu Lys
                    380                     385                     390
Pro Val Ile Val Phe Ala Asp Thr Asp Glu Asp Pro Asn Leu Leu
                    395                     400                     405
Lys Gly Ser Ala Arg Thr Thr Glu Arg Ala Leu Ala Lys Gly Tyr
                    410                     415                     420
Asn Leu Gly Glu Ala Leu Arg Lys Ala Ala Glu Leu Val Asn Gly
                    425                     430                     435
Glu Gly Gly Gly His Ala Ile Ala Ala Gly Ile Arg Ile Pro Arg
                    440                     445                     450
Ala Arg Leu Ala Glu Phe Arg Lys Leu Ile Asp Lys Ile Leu Gly
                    455                     460                     465
Glu Gln Val Ser Lys Gly Gly Asp Lys Ser Glu Ser
                    470                     475
```

SEQ ID NO: 20

SEQUENCE LENGTH: 31

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULAR TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

```
TTCATTTGGT GTTTTCCATG GATAAGGAGG G          31
```

SEQ ID NO: 21

SEQUENCE LENGTH: 23

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULAR TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

```
AAAGTWYTAA TWCCWGTNAA RGT          23
```

SEQ ID NO: 22
SEQUENCE LENGTH: 23
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
AAAGTWYTAA AAAARTTYGA RAA 23

SEQ ID NO: 23
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GATACTGCTA GAAGATTGGA 20

SEQ ID NO: 24
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TTCGTACAGT CCCTCTGGTA 20

SEQ ID NO: 25
SEQUENCE LENGTH: 957
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:

CTACGAAGCT AAAATTTGAT GTCTCAACTC AAGGACTTTT AGCTTATAAA ATGTGTCAAG 60
TCTTCCCCGA ACTTTCTCCT CCAGTAAGGG TTTTGTACCT CTCAGCAAAG ACAGGAGTAG 120
GATTTGAAGA CCTTGAAACT TTAGCGTATG AACATTATTG TACATGCGGC GACCTCACTT 180
AGATTTTTTA ACCCCTATTT TCTCTAATGT CATTCAAGTA TTGGGGGAGT AATCATGTTC 240
ACGGGTAAGG TATTGATTCC AGTAAAAGTA CTCAAGAAGT TTGAGAATTG GAATGAAGGA 300
GATATGATAC TGCTAGAAGA TTGGAAAGCC AAGGAATTGT GGGAGAGTGG AGTAGTTGAA 360
ATAATCGATG AAGCTGATAA AGTCATAGGA GAGATCGATA GAGTGTTATC AGAAGAAAAG 420
AAAAACCTCC CATTGACTCC AATACCAGAG GGACTGTACG AAAAAGCTGA ATTTTACATC 480
TATTATCTAG AAAAGTACAT CCAAGAGAAG GTCGACAACA TAGAAACAAT ACAAACTAAG 540
GTCACAAAGT TAGCAAATCT AAAGAAGAAG TATAAGACTC TGAAAGAGAT AAGATTTAAA 600
AAGATACTAG AGGCTGTGAG GCTTAGACCA AACAGTATGG AAATTCTAGC GAGATTATCC 660
CCAGCTGAAA AGAGAATATA CCTTGAGATC TCTAAAATAA GGAGAGAGTG GATAGGTGAT 720
TAGCGTGGAC AGGGAGGAGA TGATTGAGAG ATTTGCAAAC TTCCTTAGGG AGTATACAGA 780
CGAAGATGGT AACCCAGTAT ACAGAGGTAA AATAACTGAT TTACTTACAA TAACACCCAA 840
GAGGTCTGTT GCAATAGACT GGATGCACCT AAATTCCTTT GACTCAGAGC TAGAGTCGAC 900
CTGCAGGCAT GCATGCAGGT CGACTCTAGA GGATCCCCGG GTACCGAGCT CGAATTC 957

SEQ ID NO: 26
SEQUENCE LENGTH: 489
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:

ATGTTCACGG GTAAGGTATT GATTCCAGTA AAAGTACTCA AGAAGTTTGA GAATTGGAAT 60
GAAGGAGATA TGATACTGCT AGAAGATTGG AAAGCCAAGG AATTGTGGGA GAGTGGAGTA 120
GTTGAAATAA TCGATGAAGC TGATAAAGTC ATAGGAGAGA TCGATAGAGT GTTATCAGAA 180
GAAAAGAAAA ACCTCCCATT GACTCCAATA CCAGAGGGAC TGTACGAAAA AGCTGAATTT 240
TACATCTATT ATCTAGAAAA GTACATCCAA GAGAAGGTCG ACAACATAGA AACAATACAA 300
ACTAAGGTCA CAAAGTTAGC AAATCTAAAG AAGAAGTATA AGACTCTGAA AGAGATAAGA 360
TTTAAAAAGA TACTAGAGGC TGTGAGGCTT AGACCAAACA GTATGGAAAT TCTAGCGAGA 420
TTATCCCCAG CTGAAAAGAG AATATACCTT GAGATCTCTA AAATAAGGAG AGAGTGGATA 480
GGTGATTAG 489

SEQ ID NO: 27
SEQUENCE LENGTH: 162
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Phe Thr Gly Lys Val Leu Ile Pro Val Lys Val Leu Lys Lys
                5                   10                  15
Phe Glu Asn Trp Asn Glu Gly Asp Met Ile Leu Leu Glu Asp Trp
                20                  25                  30
Lys Ala Lys Glu Leu Trp Glu Ser Gly Val Val Glu Ile Ile Asp
                35                  40                  45
Glu Ala Asp Lys Val Ile Gly Glu Ile Asp Arg Val Leu Ser Glu
                50                  55                  60
Glu Lys Lys Asn Leu Pro Leu Thr Pro Ile Pro Glu Gly Leu Tyr
                65                  70                  75
Glu Lys Ala Glu Phe Tyr Ile Tyr Tyr Leu Glu Lys Tyr Ile Gln
                80                  85                  90
Glu Lys Val Asp Asn Ile Glu Thr Ile Gln Thr Lys Val Thr Lys
                95                  100                 105
Leu Ala Asn Leu Lys Lys Lys Tyr Lys Thr Leu Lys Glu Ile Arg
                110                 115                 120
Phe Lys Lys Ile Leu Glu Ala Val Arg Leu Arg Pro Asn Ser Met
                125                 130                 135
Glu Ile Leu Ala Arg Leu Ser Pro Ala Glu Lys Arg Ile Tyr Leu
                140                 145                 150
Glu Ile Ser Lys Ile Arg Arg Glu Trp Ile Gly Asp
                155                 160
```

SEQ ID NO: 28
SEQUENCE LENGTH: 17
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
ATGAAAGAYT AYAGRCC 17

SEQ ID NO: 29
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CAAGCWATWA ARGTNAAGGG 20

SEQ ID NO: 30
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TTCAAGTAAG AGTGAGTTAG 20

SEQ ID NO: 31
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TAAGTACTCC ACCATTTCCC 20

SEQ ID NO: 32
SEQUENCE LENGTH: 1012

SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:

```
TCTAGAACAT AGCAGTAAAA CTTTCCTTCT AGTACAACTT CTTCTCCTCT GTAAACTTTC   60
ACATCAACTA TCTTCTTTCT CCCTTGATCC TCCACCACCT GAGCTTTTGC TAAAAGCACG  120
TCTCCAACTT TCACCGGCTT TGTAAAGCGT ACCTCTGCCT TTCCAAGAAC TACAGTAGGC  180
TCATTTACAG CAAGCATTGC GGCGTAATCA GCTAAACCAA ATGTAAAGCC CCCGTGAACT  240
AGCCCCTTCT CATCAACCTT CATCTCGTCA ATGGTTTTCA GTTCCACTTC AGCATACCCC  300
TCTCTTATTA CCCTGGGTTT TCCTACAAGT CTCTCAGATG TCAGATTGTG CGTTTTCTGC  360
TCCATACCAC CACCGAAAAG AATAAGGTTT TTGAAATTTA AAAGCTAAGG GAGGAGTGAT  420
GAAAGACTAT AGGCCACTCC TCCAAGCAAT AAAAGTTAAG GGAGATAATG TTTTTTCAAG  480
TAAGAGTGAG TTAGTTGGTA TTCTAGCCTT TAATTTGGGA ATATTAACAG TTGGTGAGGC  540
AAAAGAACTC ATAGAGGAGG CCATAAAGGA GGGAATCATT GAGGAAACTC CCGAAGGTCT  600
CATAGTTCAT GAGGATGCCA TAACTGAAAA GGAAAGCAAA AGGGATATAT TCGGGGAAAT  660
GGTGGAGTAC TTAGCGAGAG AACTTGAGCT TAGCGAGATA GAAGTTCTTG AAGAGATAGA  720
AAAAATGAAA GAGAGGTACG GAAATTTGGA TAAAAAAATT CTTGCTTACT TATTCGGACT  780
ATCAAAAGGA GTTAACATGG AGAAATTCAA AGAATACTTG GAGGATGAAT GATGCCCAAA  840
ATAGAACCTT TTGAAAAGTA CACTGAGAGA TACGAGGAGT GGTTTGAAAG AATAAATTTG  900
CATACCTCAG TGAGCTTAAT GCCCTGAAAT CTCTTCTTCC TACCAGAGAA TGTGTTGAAG  960
TGGGAATAGG TAGTGGAAGG TTTGCGGCTC CCCTGGGAAT TAAGATGGGG GT          1012
```

SEQ ID NO: 33
SEQUENCE LENGTH: 414
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:

```
ATGAAAGACT ATAGGCCACT CCTCCAAGCA ATAAAAGTTA AGGGAGATAA TGTTTTTTCA   60
AGTAAGAGTG AGTTAGTTGG TATTCTAGCC TTTAATTTGG GAATATTAAC AGTTGGTGAG  120
GCAAAAGAAC TCATAGAGGA GGCCATAAAG GAGGGAATCA TTGAGGAAAC TCCCGAAGGT  180
CTCATAGTTC ATGAGGATGC CATAACTGAA AAGGAAAGCA AAAGGGATAT ATTCGGGGAA  240
```

ATGGTGGAGT ACTTAGCGAG AGAACTTGAG CTTAGCGAGA TAGAAGTTCT TGAAGAGATA 300
GAAAAAATGA AAGAGAGGTA CGGAAATTTG GATAAAAAAA TTCTTGCTTA CTTATTCGGA 360
CTATCAAAAG GAGTTAACAT GGAGAAATTC AAAGAATACT TGGAGGATGA ATGA 414

SEQ ID NO: 34
SEQUENCE LENGTH: 137
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Lys Asp Tyr Arg Pro Leu Leu Gln Ala Ile Lys Val Lys Gly
                  5                  10                  15
Asp Asn Val Phe Ser Ser Lys Ser Glu Leu Val Gly Ile Leu Ala
                 20                  25                  30
Phe Asn Leu Gly Ile Leu Thr Val Gly Glu Ala Lys Glu Leu Ile
                 35                  40                  45
Glu Glu Ala Ile Lys Glu Gly Ile Ile Glu Glu Thr Pro Glu Gly
                 50                  55                  60
Leu Ile Val His Glu Asp Ala Ile Thr Glu Lys Glu Ser Lys Arg
                 65                  70                  75
Asp Ile Phe Gly Glu Met Val Glu Tyr Leu Ala Arg Glu Leu Glu
                 80                  85                  90
Leu Ser Glu Ile Glu Val Leu Glu Glu Ile Glu Lys Met Lys Glu
                 95                 100                 105
Arg Tyr Gly Asn Leu Asp Lys Lys Ile Leu Ala Tyr Leu Phe Gly
                110                 115                 120
Leu Ser Lys Gly Val Asn Met Glu Lys Phe Lys Glu Tyr Leu Glu
                125                 130                 135
Asp Glu
```

SEQ ID NO: 35
SEQUENCE LENGTH: 33
SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
AAAGCTAAGG GAGGACATAT GAAAGACTAT AGG 33

SEQ ID NO: 36
SEQUENCE LENGTH: 35
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TCAAACCACT CCTCGAATTC CTCAGTGTAC TTTTC 35

SEQ ID NO: 37
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CCWTTYGARA TWGTWTTYGA 20

SEQ ID NO: 38
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GGWGCWAARG ARTTYGCNCA 20

SEQ ID NO: 39

SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
AACTTATAGA CACCGCAAGT 20

SEQ ID NO: 40
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GTCACTCTTC AACTCTTGGA 20

SEQ ID NO: 41
SEQUENCE LENGTH: 989
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:
AAGCTTATAA AAGAATACCC GATACAGACA ATGGAAAAAC TTATTTATTG AGGGGTAAAG 60
AAAGAGTTAG GCTTATGCTA AACATTCTTA AGGAGGTGGA AAGAGATGCC ATTTGAAATC 120
GTATTTGAAG GTGCAAAAGA GTTTGCCCAA CTTATAGACA CCGCAAGTAA GTTAATAGAT 180
GAGGCCGCGT TTAAAGTTAC AGAAGATGGG ATAAGCATGA GGGCCATGGA TCCAAGTAGA 240
GTTGTCCTGA TTGACCTAAA TCTCCCGTCA AGCATATTTA GCAAATATGA AGTTGTTGAA 300
CCAGAAACAA TTGGAGTTAA CATGGACCAC CTAAAGAAGA TCCTAAAGAG AGGTAAAGCA 360
AAGGACACCT TAATACTCAA GAAAGGAGAG GAAAACTTCT TAGAGATAAC AATTCAAGGA 420
ACTGCAACAA GAACATTTAG AGTTCCCCTA ATAGATGTAG AAGAGATGGA AGTTGACCTC 480
CCAGAACTTC CATTCACTGC AAAGGTTGTA GTTCTTGGAG AAGTCCTAAA AGATGCTGTT 540
AAAGATGCCT CTCTAGTGAG TGACAGCATA AAATTTATTG CCAGGGAAAA TGAATTTATA 600

ATGAAGGCAG AGGGAGAAAC CCAGGAAGTT GAGATAAAGC TAACTCTTGA AGATGAGGGA 660
TTATTGGACA TCGAGGTTCA AGAGGAGACA AAGAGCGCAT ATGGAGTCAG CTATCTCTCC 720
GACATGGTTA AAGGACTTGG AAAGGCCGAT GAAGTTACAA TAAAGTTTGG AAATGAAATG 780
CCCATGCAAA TGGAGTATTA CATTAGAGAT GAAGGAAGAC TTACATTCCT ACTGGCTCCA 840
AGAGTTGAAG AGTGACTTTT CTTTTCCTTA TAATTTAATT TGGGGATAAC AATGGATATT 900
GAGGTTCTCA GAAGATTATT GGAGAGAGAA CTTTCAAGCG AAGAACTGAC TAAAATAGAG 960
GAAGAATTTT ATGACGATTT AGAAAGCTT 989

SEQ ID NO: 42
SEQUENCE LENGTH: 45
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CCGGAACCGC CTCCCTCAGA GCCGCCACCC TCAGAACCGC CACCC 45

SEQ ID NO: 43
SEQUENCE LENGTH: 23
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CCWTGGGTWG ARAARTAYAG RCC 23

SEQ ID NO: 44
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
WSWGATGAAA GAGGNATHGA 20

SEQ ID NO: 45
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GCWTTWAGAA GAACNATGGA 20

SEQ ID NO: 46
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TTWCCWACWC CWGGWGGNCC 20

SEQ ID NO: 47
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CTTCTTAAWG CATTYTGNGC 20

SEQ ID NO: 48
SEQUENCE LENGTH: 23
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:

ATWATTTTWS WWGGATARTT RCA 23

SEQ ID NO: 49
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
ATWGCTTTTC TCATRTCNCC 20

SEQ ID NO: 50
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
ATCTTGAGTT AAAGCGTCGG 20

SEQ ID NO: 51
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
ACGTTCGCTT TATCTTGAGC 20

SEQ ID NO: 52
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TCAAAGACTT GACGACATTG 20

SEQ ID NO: 53
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TTCTGCTATG TAAAGTATTG 20

SEQ ID NO: 54
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CAATACTTTA CATAGCAGAA 20

SEQ ID NO: 55
SEQUENCE LENGTH: 3620
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:
GAGCTCCAGC AACAACAATA ACCCAAGATG GAAAGGACTT TGGAGTAAGG TACTTTGGAT 60
TACCGGCAGG TCATGAGTTC GCAGCATTCT TAGAGGACAT TGTGGATGTT AGTAGAGAAG 120
AAACAAACCT TATGGACGAG ACAAAACAGG CCATCAGAAA CATAGACCAG GATGTAAGAA 180
TATTGGTGTT TGAAACTCCA ACATGCCCAT ACTGTCCACT TGCCGTTAGA ATGGCTCACA 240
AGTTTGCCAT TGAAAACACA AAAGCTGGGA AAGGTAAGAT ACTTGGGGAT ATGGTCGAGG 300

CCATTGAGTA TCCAGAGTGG GCTGACCAGT ACAATGTAAT GGCAGTACCA AAAATTGTTA 360
TTCAGGTCAA CGGAGAAGAC AGAGTAGAAT TTGAAGGAGC TTATCCAGAG AAAATGTTCT 420
TAGAGAAGTT ACTCTCAGCT CTCAGCTGAT CTACTGTTTT TCCTTCTTTT CTTCTGTTCT 480
GTTATTGCCT AGGATAAGCT TAATAATACT TTGATACCTT TCTTAGTTTA GGTGTGTGAG 540
AGTATGAGCG AAGAGATTAG AGAAGTTAAG GTTCTAGAAA AACCCTGGGT TGAGAAGTAT 600
AGACCTCAAA GACTTGACGA CATTGTAGGA CAAGAGCACA TAGTGAAAAG GCTCAAGCAC 660
TACGTCAAAA CTGGATCAAT GCCCCACCTA CTCTTCGCAG GCCCCCCTGG TGTCGGAAAG 720
TGTCTTACTG GAGATACCAA AGTTATAGCT AATGGCCAAC TCTTTGAACT TGGAGAACTT 780
GTTGAAAAGC TTTCTGGGGG GAGATTTGGA CCAACTCCAG TTAAAGGGCT CAAAGTTCTT 840
GGAATAGATG AGGATGGAAA GCTTAGAGAG TTTGAAGTCC AATACGTCTA CAAAGATAGA 900
ACTGATAGGT TGATAAAGAT AAAAACTCAC CTTGGCAGGG AGCTTAAAGT AACTCCGTAT 960
CACCCACTTC TAGTGAATAG AGAGAATGGC GAAATAAAGT GGATTAAGGC TGAAGAACTC 1020
AAACCTGGCG ACAAGCTTGC AATACCGAGC TTTCTCCCAC TTATAACTGG AGAAAATCCC 1080
CTTGCAGAGT GGCTTGGTTA CTTTATGGGA AGTGGCTATG CTTATCCAAG TAATTCTGTC 1140
ATCACGTTCA CTAACGAAGA TCCACTCATA AGACAACGCT TTATGGAACT AACAGAGAAA 1200
CTTTTCCCTG ATGCAAAGAT AAGGGAAAGA ATTCACGCTG ATGGAACTCC AGAAGTTTAT 1260
GTGGTATCTA GGAAAGCTTG GAGCCTTGTA AACTCTATTA GCTTAACATT AATACCCAGG 1320
GAGGGGTGGA AAGGAATTCG TTCTTTCCTT AGGGCATATT CCGACTGCAA TGGTCGGATT 1380
GAAAGTGATG CAATAGTTTT ATCAACCGAT AACAATGATA TGGCCCAGCA GATAGCCTAT 1440
GCTTTAGCCA GCTTTGGAAT AATAGCTAAA ATGGATGGAG AAGATGTTAT TATCTCAGGC 1500
TCGGACAACA TAGAGAGGTT CCTAAATGAG ATTGGCTTTA GCACCCAAAG CAAACTTAAA 1560
GAAGCCCAGA AGCTCATTAG AAAAACCAAT GTAAGATCCG ATGGACTAAA GATTAACTAT 1620
GAGCTAATCT CCTATGTAAA AGACAGGCTT AGGTTAAATG TCAATGATAA AAGAAATTTG 1680
AGCTACAGAA ATGCAAAGGA GCTTTCTTGG GAACTCATGA AAGAAATTTA TTATCGCCTT 1740
GAGGAACTGG AGAGACTAAA GAAGGTCTTA TCAGAACCCA TCTTGATCGA CTGGAATGAA 1800
GTAGCAAAGA AGAGTGATGA AGTAATAGAA AAAGCTAAAA TTAGAGCAGA GAAGCTCCTA 1860
GAATACATAA AAGGAGAGAG AAAGCCAAGT TTCAAGGAGT ACATTGAGAT AGCAAAAGTC 1920
CTTGGAATTA ACGTTGAACG TACCATCGAA GCTATGAAGA TCTTTGCAAA GAGATACTCA 1980
AGCTATGCCG AGATTGGAAG AAAACTTGGA ACTTGGAATT TCAATGTAAA AACAATTCTT 2040
GAGAGCGACA CAGTGGATAA CGTTGAAATC CTTGAAAAGA TAAGGAAAAT TGAGCTTGAG 2100
CTCATAGAGG AAATTCTTTC GGATGGAAAG CTCAAAGAAG GTATAGCATA TCTCATTTTC 2160
CTCTTCCAGA ATGAGCTTTA CTGGGACGAG ATAACTGAAG TAAAAGAGCT TAGGGGAGAC 2220
TTTATAATCT ATGATCTTCA TGTTCCTGGC TACCACAACT TTATTGCTGG GAACATGCCA 2280
ACAGTAGTCC ATAACACTAC AGCGGCTTTG GCCCTTGCAA GAGAGCTTTT CGGCGAAAAC 2340

TGGAGGCATA ACTTCCTCGA GTTGAATGCT TCAGATGAAA GAGGTATAAA CGTAATTAGA 2400
GAGAAAGTTA AGGAGTTTGC GAGAACAAAG CCTATAGGAG GAGCAAGCTT CAAGATAATT 2460
TTCCTTGATG AGGCCGACGC TTTAACTCAA GATGCCCAAC AAGCCTTAAG AAGAACCATG 2520
GAAATGTTCT CGAGTAACGT TCGCTTTATC TTGAGCTGTA ACTACTCCTC CAAGATAATT 2580
GAACCCATAC AGTCTAGATG TGCAATATTC CGCTTCAGAC CTCTCCGCGA TGAGGATATA 2640
GCGAAGAGAC TAAGGTACAT TGCCGAAAAT GAGGGCTTAG AGCTAACTGA AGAAGGTCTC 2700
CAAGCAATAC TTTACATAGC AGAAGGAGAT ATGAGAAGAG CAATAAACAT TCTGCAAGCT 2760
GCAGCAGCTC TAGACAAGAA GATCACCGAC GAAAACGTAT TCATGGTAGC GAGTAGAGCT 2820
AGACCTGAAG ATATAAGAGA GATGATGCTT CTTGCTCTCA AAGGCAACTT CTTGAAGGCC 2880
AGAGAAAAGC TTAGGGAGAT ACTTCTCAAG CAAGGACTTA GTGGAGAAGA TGTACTAGTT 2940
CAGATGCACA AAGAAGTCTT CAACCTGCCA ATAGAGGAGC CAAAGAAGGT TCTGCTTGCT 3000
GATAAGATAG GAGAGTATAA CTTCAGACTC GTTGAAGGGG CTAATGAAAT AATTCAGCTT 3060
GAAGCACTCT TAGCACAGTT CACCCTAATT GGGAAGAAGT GATGAAGTAT GCCAGAGCTT 3120
NCCTTGGGTA GAAAAATACA GGCCAAAAAA GCTAAGTGAA ATTGTAAACC AAGAAGAGGC 3180
TATAGAGAAA GTTAGAGCGT GGATAGAGAG CTGGTTGCAT GNCCACCCCC TTNAGAAAAA 3240
AGCCGTATTA TTAGCAGGAC CCCCAGGGAG CGGAAAGACA ACCACAGTNT ACGCTNTAGC 3300
AAATGAGTAC AACTTTGAAG TCATTGAGCT CAACGCGAGT GATGAGAGAA CTTATGAAAA 3360
AATCTCCAGG TATGTTCAAG CAGCATACAC TATGGATATC CTCGGAAAGA GGAGGAAGAT 3420
AATCTTCCTC GATGAAGCAG ATAATATAGA GCCCAGCGGA GCTAAGGAAA TCGCAAAACT 3480
AATTGATAAG GCCAAAAATC CAATAATAAT GGCTGCAAAT AAGTACTGGG AAGTTCCAAA 3540
AGAGATCCGA GAAAAAGCTG AGCTAGTAGA GTACAAGAGG TTAACCCAGA GAGATGTAAT 3600
GAATGCCTTA ATAAGGATCC 3620

SEQ ID NO: 56
SEQUENCE LENGTH: 21
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CTTTCCGACA CCAGGGGGGC C 21

SEQ ID NO: 57
SEQUENCE LENGTH: 21

SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
ACTACAGCGG CTTTGGCCCT T 21

SEQ ID NO: 58
SEQUENCE LENGTH: 23
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GATGAGTTCG TGTCCGTACA ACT 23

SEQ ID NO: 59
SEQUENCE LENGTH: 22
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
ACAAAGCCAG CCGGAATATC TG 22

SEQ ID NO: 60
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GCTTCTAAAT CATTDATNGC 20

SEQ ID NO: 61
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GCGTGGATAG AGAGCTGGTT 20

SEQ ID NO: 62
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CTCTGGGTTA ACCTCTTGTA 20

SEQ ID NO: 63
SEQUENCE LENGTH: 1437
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:
ATGCCAGAGC TTCCCTGGGT AGAAAAATAC AGGCCAAAAA AGTTAAGTGA AATTGTAAAC 60
CAAGAAGAGG CTATAGAGAA AGTTAGAGCG TGGATAGAGA GCTGGTTGCA TGGCCACCCC 120
CCTAAGAAAA AAGCCCTATT ATTAGCAGGA CCCCCAGGGA GCGGAAAGAC AACCACAGTC 180
TACGCTCTAG CAAATGAGTA CAACTTTGAA GTCATTGAGC TCAACGCGAG TGATGAGAGA 240
ACTTATGAAA AAATCTCCAG GTATGTTCAA GCAGCATACA CTATGGATAT CCTCGGAAAG 300
AGGAGGAAGA TAATCTTCCT CGATGAAGCA GATAATATAG AGCCCAGCGG AGCTAAGGAA 360
ATCGCAAAAC TAATTGATAA GGCCAAAAAT CCAATAATAA TGGCTGCAAA TAAGTACTGG 420
GAAGTTCCAA AAGAGATCCG AGAAAAAGCT GAGCTAGTAG AGTACAAGAG GTTAACCCAG 480
AGAGATGTAA TGAATGCCTT AATAAGGATC CTAAAGAGGG AAGGTATAAC AGTTCCAAAA 540

GAAATCCTCC TAGAAATAGC AAAAAGATCT AGTGGAGATC TAAGAGCAGC TATAAATGAT 600
CTACAGACCG TTGTAGTGGG TGGTTACGAA GATGCTACGC AAGTTTTGGC ATATAGAGAT 660
GTAGAAAAGA CAGTCTTTCA AGCCCTAGGA CTCGTCTTTG GAAGTGACAA CGCCAAGAGG 720
GCAAAGATGG CAATGTGGAA CTTGGACATG TCCCCTGATG AATTCCTGCT ATGGGTAGAT 780
GAGAACATTC CTCACCTCTA CCTAAATCCA GAGGAGATTG CCCAGGCGTA TGATGCAATT 840
AGTAGAGCCG ACATATACCT CGGAAGGGCC GCCAGAACTG GAAACTATTC ACTCTGGAAG 900
TACGCAATAG ATATGATGAC TGCAGGAGTT GCCGTGGCAG GGAGAAAGAG AAGGGGATTT 960
GTCAAGTTTT ATCCTCCCAA CACCCTAAAG ATTTTAGCGG AAAGCAAAGA AGAAAGAGAG 1020
ATCAGAGAGT CAATAATTAA AAAGATAATA CGAGAGATGC ACATGAGTAG GCTACAGGCA 1080
ATAGAAACGA TGAAAATAAT TAGAGAGATT TTCGAGAACA ATCTAGACCT TGCTGCGCAC 1140
TTTACAGTGT TCCTTGGTCT GTCTGAAAAA GAAGTTGAGT TTCTAGCTGG AAAGGAAAAA 1200
GCTGGTACCA TTTGGGGCAA AGCCTTAGCA TTAAGAAGGA AACTTAAGGA GCTTGGAATA 1260
AGAGAGGAGG AGAAGCCTAA AGTTGAAATT GAAGAAGAGG AAGAAGAGGA AGAAAAGACC 1320
GAAGAAGAAA AAGAGGAAAT AGAAGAAAAA CCCGAAGAAG AGAAAGAAGA GGAGAAGAAA 1380
GAAAAGGAAA AGCCAAAGAA AGGCAAACAA GCAACTCTCT TTGACTTTCT TAAAAAG 1437

SEQ ID NO: 64
SEQUENCE LENGTH: 479
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

Met Pro Glu Leu Pro Trp Val Glu Lys Tyr Arg Pro Lys Lys Leu
5 10 15
Ser Glu Ile Val Asn Gln Glu Glu Ala Ile Glu Lys Val Arg Ala
20 25 30
Trp Ile Glu Ser Trp Leu His Gly His Pro Pro Lys Lys Lys Ala
35 40 45
Leu Leu Leu Ala Gly Pro Pro Gly Ser Gly Lys Thr Thr Thr Val
50 55 60
Tyr Ala Leu Ala Asn Glu Tyr Asn Phe Glu Val Ile Glu Leu Asn
65 70 75
Ala Ser Asp Glu Arg Thr Tyr Glu Lys Ile Ser Arg Tyr Val Gln

80 85 90
Ala Ala Tyr Thr Met Asp Ile Leu Gly Lys Arg Arg Lys Ile Ile
95 100 105
Phe Leu Asp Glu Ala Asp Asn Ile Glu Pro Ser Gly Ala Lys Glu
110 115 120
Ile Ala Lys Leu Ile Asp Lys Ala Lys Asn Pro Ile Ile Met Ala
125 130 135
Ala Asn Lys Tyr Trp Glu Val Pro Lys Glu Ile Arg Glu Lys Ala
140 145 150
Glu Leu Val Glu Tyr Lys Arg Leu Thr Gln Arg Asp Val Met Asn
155 160 165
Ala Leu Ile Arg Ile Leu Lys Arg Glu Gly Ile Thr Val Pro Lys
170 175 180
Glu Ile Leu Leu Glu Ile Ala Lys Arg Ser Ser Gly Asp Leu Arg
185 190 195
Ala Ala Ile Asn Asp Leu Gln Thr Val Val Val Gly Gly Tyr Glu
200 205 210
Asp Ala Thr Gln Val Leu Ala Tyr Arg Asp Val Glu Lys Thr Val
215 220 225
Phe Gln Ala Leu Gly Leu Val Phe Gly Ser Asp Asn Ala Lys Arg
230 235 240
Ala Lys Met Ala Met Trp Asn Leu Asp Met Ser Pro Asp Glu Phe
245 250 255
Leu Leu Trp Val Asp Glu Asn Ile Pro His Leu Tyr Leu Asn Pro
260 265 270
Glu Glu Ile Ala Gln Ala Tyr Asp Ala Ile Ser Arg Ala Asp Ile
275 280 285
Tyr Leu Gly Arg Ala Ala Arg Thr Gly Asn Tyr Ser Leu Trp Lys
290 295 300
Tyr Ala Ile Asp Met Met Thr Ala Gly Val Ala Val Ala Gly Arg
305 310 315
Lys Arg Arg Gly Phe Val Lys Phe Tyr Pro Pro Asn Thr Leu Lys
320 325 330
Ile Leu Ala Glu Ser Lys Glu Glu Arg Glu Ile Arg Glu Ser Ile

335 340 345
Ile Lys Lys Ile Ile Arg Glu Met His Met Ser Arg Leu Gln Ala
350 355 360
Ile Glu Thr Met Lys Ile Ile Arg Glu Ile Phe Glu Asn Asn Leu
365 370 375
Asp Leu Ala Ala His Phe Thr Val Phe Leu Gly Leu Ser Glu Lys
380 385 390
Glu Val Glu Phe Leu Ala Gly Lys Glu Lys Ala Gly Thr Ile Trp
395 400 405
Gly Lys Ala Leu Ala Leu Arg Arg Lys Leu Lys Glu Leu Gly Ile
410 415 420
Arg Glu Glu Glu Lys Pro Lys Val Glu Ile Glu Glu Glu Glu Glu
425 430 435
Glu Glu Glu Lys Thr Glu Glu Glu Lys Glu Glu Ile Glu Glu Lys
440 445 450
Pro Glu Glu Glu Lys Glu Glu Glu Lys Lys Glu Lys Glu Lys Pro
455 460 465
Lys Lys Gly Lys Gln Ala Thr Leu Phe Asp Phe Leu Lys Lys
470 475

SEQ ID NO: 65
SEQUENCE LENGTH: 23
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
ATGGATATWG ARGTDYTNAG RAG 23

SEQ ID NO: 66
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
ATWGARGTWY TWAGRAGRYT 20

SEQ ID NO: 67
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GAGAGAGAAC TTTCAAGCGA 20

SEQ ID NO: 68
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CTCTAAGAAG ATATGCCTCT 20

SEQ ID NO: 69
SEQUENCE LENGTH: 558
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:
ATGGATATTG AGGTTCTCAG AAGATTATTG GAGAGAGAAC TTTCAAGCGA AGAACTGACT 60
AAAATAGAGG AAGAATTTTA TGACGATTTA GAAAGCTTTA GAAAAGCCTT GGAAATCAAT 120
GCCGAGAGAC ATGAAGAAAG AGGAGAGGAC ATTCACAAAA AGCTGTATTT AGCTCAACTA 180
TCTTTGGTTA GGAATCTTGT TAGAGAAATA TTAAGGATTA GGTTGCATAA GATTGTTGAT 240
ATGGCATTTG AGGGAGTTCC CAGAAATTTA GTTGGAGATG AAAAGAAAAT ATACAAGATA 300

```
ATAACAGCTT TCATAAATGG AGAACCTCTT GAAATTGAAA CGGCAGGAGA AGAGAGTATT  360
GAAGTTATTG AAGAGGAAAA AGAAACATCT CCTGGGATAA TAGAGGCATA TCTTCTTAGA  420
GTTGATATTC CCAAAATATT GGATGAAAAT TTGAGAGAAT ATGGGCCCTT CAAGGCTGGC  480
GATCTTGTTG TATTGCCGAA GTCTATTGGC AGGGTACTCA TTCAGAGGGA TGCCGCGGAT  540
AAGGTATTGA TACAATTG                                                558
```

SEQ ID NO: 70
SEQUENCE LENGTH: 186
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Asp Ile Glu Val Leu Arg Arg Leu Leu Glu Arg Glu Leu Ser
                5                   10                  15
Ser Glu Glu Leu Thr Lys Ile Glu Glu Glu Phe Tyr Asp Asp Leu
                20                  25                  30
Glu Ser Phe Arg Lys Ala Leu Glu Ile Asn Ala Glu Arg His Glu
                35                  40                  45
Glu Arg Gly Glu Asp Ile His Lys Lys Leu Tyr Leu Ala Gln Leu
                50                  55                  60
Ser Leu Val Arg Asn Leu Val Arg Glu Ile Leu Arg Ile Arg Leu
                65                  70                  75
His Lys Ile Val Asp Met Ala Phe Glu Gly Val Pro Arg Asn Leu
                80                  85                  90
Val Gly Asp Glu Lys Lys Ile Tyr Lys Ile Ile Thr Ala Phe Ile
                95                  100                 105
Asn Gly Glu Pro Leu Glu Ile Glu Thr Ala Gly Glu Glu Ser Ile
                110                 115                 120
Glu Val Ile Glu Glu Glu Lys Glu Thr Ser Pro Gly Ile Ile Glu
                125                 130                 135
Ala Tyr Leu Leu Arg Val Asp Ile Pro Lys Ile Leu Asp Glu Asn
                140                 145                 150
Leu Arg Glu Tyr Gly Pro Phe Lys Ala Gly Asp Leu Val Val Leu
```

155 160 165
Pro Lys Ser Ile Gly Arg Val Leu Ile Gln Arg Asp Ala Ala Asp
170 175 180
Lys Val Leu Ile Gln Leu
185

SEQ ID NO: 71
SEQUENCE LENGTH: 33
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TTTAATTTGG GGATAACCAT GGATATTGAG GTT 33

SEQ ID NO: 72
SEQUENCE LENGTH: 31
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TAGGATGGGT TTTGGATCCT CTCATTGGAG G 31

SEQ ID NO: 73
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
ATGATWGGWW SWATHTTYTA 20

SEQ ID NO: 74

SEQUENCE LENGTH: 23
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
AAGAAGTTTA ATYTDCAYAG RCC 23

SEQ ID NO: 75
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TGAGTATCAT CCAGAGAATC 20

SEQ ID NO: 76
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TCACATCGGG ATCGTTCCAG 20

SEQ ID NO: 77
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GATTTTGACG CTCATCATGG 20

SEQ ID NO: 78
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GGAAAGAACG ATTTCGAGTC 20

SEQ ID NO: 79
SEQUENCE LENGTH: 1005
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: Genomic DNA
SEQUENCE DESCRIPTION:
ATGATTGGCT CAATATTTTA TTCCAAGAAG TTTAACCTCC ATAGACCTAG TGAGTATCAT 60
CCAGAGAATC CCAAGAGACT CGAAATCGTT CTTTCCAAGG TCAGAGAGCT TGGACTTGAA 120
GAAAGAATAG AAGAACCAAA CCCAGTTGAA GAGACTTTCG TTGAGAAAAT TCACGACAGG 180
GATTACATCA ACTTCGTTAA AGAGGCCGTT GAAAAAGGAA TCACAAGACT TGATCCAGAC 240
ACTTATGTTT CTCCTGGGAC TTGGAGTGCG GCATTGTTAG CTTTAGGAGC CGCAAGGAGT 300
GCAGCTTTAT CAGCCCTTCA CTATGGAGGC CTCCACATGG CTCTAGTTAG GCCCCCTGGG 360
CATCATGCAG GGAGAAGAGG AAGGGCCATG GGTGCCCCAA CACTAGGCTT CTGCATCTTC 420
AACAACGCGG CCTCTGCAGT TGTCACCTTG AAAGAAGAGG GAGTTGGAAA AGTTGTTGTA 480
ATAGATTTTG ACGCTCATCA TGGAAACGGG ACTCAGGAAA TATTCTGGAA CGATCCCGAT 540
GTGATTCACA TAGATCTACA CGAGAGAGAC ATCTACCCAG GGAGTGGGGA TGTGAGTGAA 600
GTTGGAGGGT CAAATGCTTA TGGGAGCAAG ATAAACCTCC CAATGCCCCA CTATTCTGGG 660
GATGGGGATT ACATATATGT TTGGGACGAA ATTGTGCTTC CAATAGTTGA AGAAGTTAAG 720
CCAAAGGTCA TCGTAATTTC CGCGGGCTTT GATGGATTTA AAGGGGATGG TCTAACAACA 780
TTAAGGCTCA CAGAAAGTTT TTACTCTTAT GCAGGGGCTA CATTAAATAA ATATCCCTTG 840
GCATTTATAT TGGAAGGCGG GTATGGAGTA GGGTTAGATA AAGGTTTTCC GGCCTTCATA 900
ATGGGCTACG AAGAGGGTAA AGCGAAAGCT CGAGAAGAGC CAAGATATGA GACCCTAAAG 960
TTGGCGGAGG AGGTTAAGGA CATCTTGAGT CCCTGGTGGT CGTTA 1005

SEQ ID NO: 80
SEQUENCE LENGTH: 335
SEQUENCE TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
SEQUENCE DESCRIPTION:

```
Met Ile Gly Ser Ile Phe Tyr Ser Lys Lys Phe Asn Leu His Arg
                  5                  10                  15
Pro Ser Glu Tyr His Pro Glu Asn Pro Lys Arg Leu Glu Ile Val
                 20                  25                  30
Leu Ser Lys Val Arg Glu Leu Gly Leu Glu Glu Arg Ile Glu Glu
                 35                  40                  45
Pro Asn Pro Val Glu Glu Thr Phe Val Glu Lys Ile His Asp Arg
                 50                  55                  60
Asp Tyr Ile Asn Phe Val Lys Glu Ala Val Glu Lys Gly Ile Thr
                 65                  70                  75
Arg Leu Asp Pro Asp Thr Tyr Val Ser Pro Gly Thr Trp Ser Ala
                 80                  85                  90
Ala Leu Leu Ala Leu Gly Ala Ala Arg Ser Ala Ala Leu Ser Ala
                 95                 100                 105
Leu His Tyr Gly Gly Leu His Met Ala Leu Val Arg Pro Pro Gly
                110                 115                 120
His His Ala Gly Arg Arg Gly Arg Ala Met Gly Ala Pro Thr Leu
                125                 130                 135
Gly Phe Cys Ile Phe Asn Asn Ala Ala Ser Ala Val Val Thr Leu
                140                 145                 150
Lys Glu Glu Gly Val Gly Lys Val Val Val Ile Asp Phe Asp Ala
                155                 160                 165
His His Gly Asn Gly Thr Gln Glu Ile Phe Trp Asn Asp Pro Asp
                170                 175                 180
Val Ile His Ile Asp Leu His Glu Arg Asp Ile Tyr Pro Gly Ser
                185                 190                 195
Gly Asp Val Ser Glu Val Gly Gly Ser Asn Ala Tyr Gly Ser Lys
```

200 205 210
Ile Asn Leu Pro Met Pro His Tyr Ser Gly Asp Gly Asp Tyr Ile
215 220 225
Tyr Val Trp Asp Glu Ile Val Leu Pro Ile Val Glu Glu Val Lys
230 235 240
Pro Lys Val Ile Val Ile Ser Ala Gly Phe Asp Gly Phe Lys Gly
245 250 255
Asp Gly Leu Thr Thr Leu Arg Leu Thr Glu Ser Phe Tyr Ser Tyr
260 265 270
Ala Gly Ala Thr Leu Asn Lys Tyr Pro Leu Ala Phe Ile Leu Glu
275 280 285
Gly Gly Tyr Gly Val Gly Leu Asp Lys Gly Phe Pro Ala Phe Ile
290 295 300
Met Gly Tyr Glu Glu Gly Lys Ala Lys Ala Arg Glu Glu Pro Arg
305 310 315
Tyr Glu Thr Leu Lys Leu Ala Glu Glu Val Lys Asp Ile Leu Ser
320 325 330
Pro Trp Trp Ser Leu
335

SEQ ID NO: 81
SEQUENCE LENGTH: 36
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GGGAAGAAGT GATGACATAT GCCAGAGCTT CCCTGG 36

SEQ ID NO: 82
SEQUENCE LENGTH: 20
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TTCCAAGCTC CTTAAGTTTC 20

SEQ ID NO: 83
SEQUENCE LENGTH: 3574
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CATATGCCAG AGCTTCCCTG GGTAGAAAAA TACAGGCCAA AAAAGTTAAG TGAAATTGTA 60
AACCAAGAAG AGGCTATAGA GAAAGTTAGA GCGTGGATAG AGAGCTGGTT GCATGGCCAC 120
CCCCCTAAGA AAAAAGCCCT ATTATTAGCA GGACCCCCAG GGAGCGGAAA GACAACCACA 180
GTCTACGCTC TAGCAAATGA GTACAACTTT GAAGTCATTG AGCTCAACGC GAGTGATGAG 240
AGAACTTATG AAAAAATCTC CAGGTATGTT CAAGCAGCAT ACACTATGGA TATCCTCGGA 300
AAGAGGAGGA AGATAATCTT CCTCGATGAA GCAGATAATA TAGAGCCCAG CGGAGCTAAG 360
GAAATCGCAA AACTAATTGA TAAGGCCAAA AATCCAATAA TAATGGCTGC AAATAAGTAC 420
TGGGAAGTTC CAAAAGAGAT CCGAGAAAAA GCTGAGCTAG TAGAGTACAA GAGGTTAACC 480
CAGAGAGATG TAATGAATGC CTTAATAAGG ATCCTAAAGA GGGAAGGTAT AACAGTTCCA 540
AAAGAAATCC TCCTAGAAAT AGCAAAAAGA TCTAGTGGAG ATCTAAGAGC AGCTATAAAT 600
GATCTACAGA CCGTTGTAGT GGGTGGTTAC GAAGATGCTA CGCAAGTTTT GGCATATAGA 660
GATGTAGAAA AGACAGTCTT TCAAGCCCTA GGACTCGTCT TTGGAAGTGA CAACGCCAAG 720
AGGGCAAAGA TGGCAATGTG GAACTTGGAC ATGTCCCCTG ATGAATTCCT GCTATGGGTA 780
GATGAGAACA TTCCTCACCT CTACCTAAAT CCAGAGGAGA TTGCCCAGGC GTATGATGCA 840
ATTAGTAGAG CCGACATATA CCTCGGAAGG GCCGCCAGAA CTGGAAACTA TTCACTCTGG 900
AAGTACGCAA TAGATATGAT GACTGCAGGA GTTGCCGTGG CAGGGAGAAA GAGAAGGGGA 960
TTTGTCAAGT TTTATCCTCC CAACACCCTA AAGATTTTAG CGGAAAGCAA AGAAGAAAGA 1020
GAGATCAGAG AGTCAATAAT TAAAAAGATA ATACGAGAGA TGCACATGAG TAGGCTACAG 1080
GCAATAGAAA CGATGAAAAT AATTAGAGAG ATTTTCGAGA ACAATCTAGA CCTTGCTGCG 1140
CACTTTACAG TGTTCCTTGG TCTGTCTGAA AAAGAAGTTG AGTTTCTAGC TGGAAAGGAA 1200
AAAGCTGGTA CCATTTGGGG CAAAGCCTTA GCATTAAGAA GGAAACTTAA GGAGCTTGGA 1260
ATAAGAGAGG AGGAGAAGCC TAAAGTTGAA ATTGAAGAAG AGGAAGAAGA GGAAGAAAAG 1320
ACCGAAGAAG AAAAAGAGGA AATAGAAGAA AAACCCGAAG AAGAGAAAGA AGAGGAGAAG 1380

AAAGAAAAGG AAAAGCCAAA GAAAGGCAAA CAAGCAACTC TCTTTGACTT TCTTAAAAAG 1440
TGATTACCCT TTTTCTTCTA TTAGAGCTCC GAATAAAGTT GGCCCTCTAA TTTTTTCTAT 1500
TGTCTCCTCC ACATTAATCT TTACGAATTC GAGCTCCAGC AACAACAATA ACCCAAGATG 1560
GAAAGGACTT TGGAGTAAGG TACTTTGGAT TACCGGCAGG TCATGAGTTC GCAGCATTCT 1620
TAGAGGACAT TGTGGATGTT AGTAGAGAAG AAACAAACCT TATGGACGAG ACAAAACAGG 1680
CCATCAGAAA CATAGACCAG GATGTAAGAA TATTGGTGTT TGAAACTCCA ACATGCCCAT 1740
ACTGTCCACT TGCCGTTAGA ATGGCTCACA AGTTTGCCAT TGAAAACACA AAAGCTGGGA 1800
AAGGTAAGAT ACTTGGGGAT ATGGTCGAGG CCATTGAGTA TCCAGAGTGG GCTGACCAGT 1860
ACAATGTAAT GGCAGTACCA AAAATTGTTA TTCAGGTCAA CGGAGAAGAC AGAGTAGAAT 1920
TTGAAGGAGC TTATCCAGAG AAAATGTTCT TAGAGAAGTT ACTCTCAGCT CTCAGCTGAT 1980
CTACTGTTTT TCCTTCTTTT CTTCTGTTCT GTTATTGCCT AGGATAAGCT TAATAATACT 2040
TTGATACCTT TCTTAGTTTA GGTGTGTGAG AGTATGAGCG AAGAGATTAG AGAAGTTAAG 2100
GTTCTAGAAA AACCCTGGGT TGAGAAGTAT AGACCTCAAA GACTTGACGA CATTGTAGGA 2160
CAAGAGCACA TAGTGAAAAG GCTCAAGCAC TACGTCAAAA CTGGATCAAT GCCCCACCTA 2220
CTCTTCGCAG GCCCCCCTGG TGTCGGAAAG ACTACAGCGG CTTTGGCCCT TGCAAGAGAG 2280
CTTTTCGGCG AAAACTGGAG GCATAACTTC CTCGAGTTGA ATGCTTCAGA TGAAAGAGGT 2340
ATAAACGTAA TTAGAGAGAA AGTTAAGGAG TTTGCGAGAA CAAAGCCTAT AGGAGGAGCA 2400
AGCTTCAAGA TAATTTTCCT TGATGAGGCC GACGCTTTAA CTCAAGATGC CCAACAAGCC 2460
TTAAGAAGAA CCATGGAAAT GTTCTCGAGT AACGTTCGCT TTATCTTGAG CTGTAACTAC 2520
TCCTCCAAGA TAATTGAACC CATACAGTCT AGATGTGCAA TATTCCGCTT CAGACCTCTC 2580
CGCGATGAGG ATATAGCGAA GAGACTAAGG TACATTGCCG AAAATGAGGG CTTAGAGCTA 2640
ACTGAAGAAG GTCTCCAAGC AATACTTTAC ATAGCAGAAG GAGATATGAG AAGAGCAATA 2700
AACATTCTGC AAGCTGCAGC AGCTCTAGAC AAGAAGATCA CCGACGAAAA CGTATTCATG 2760
GTAGCGAGTA GAGCTAGACC TGAAGATATA AGAGAGATGA TGCTTCTTGC TCTCAAAGGC 2820
AACTTCTTGA AGGCCAGAGA AAAGCTTAGG GAGATACTTC TCAAGCAAGG ACTTAGTGGA 2880
GAAGATGTAC TAGTTCAGAT GCACAAAGAA GTCTTCAACC TGCCAATAGA GGAGCCAAAG 2940
AAGGTTCTGC TTGCTGATAA GATAGGAGAG TATAACTTCA GACTCGTTGA AGGGGCTAAT 3000
GAAATAATTC AGCTTGAAGC ACTCTTAGCA CAGTTCACCC TAATTGGGAA GAAGTGATGA 3060
AGTATGCCAG AGCTTCCCTG GGTAGAAAAA TACAGGCCAA AAAAGTTAAG TGAAATTGTA 3120
AACCAAGAAG AGGCTATAGA GAAAGTTAGA GCGTGGATAG AGAGCTGGTT GCATGGCCAC 3180
CCCCCTAAGA AAAAAGCCGT ATTATTAGCA GGACCCCCAG GGAGCGGAAA GACAACCACA 3240
GTCTACGCTC TAGCAAATGA GTACAACTTT GAAGTCATTG AGCTCAACGC GAGTGATGAG 3300
AGAACTTATG AAAAAATCTC CAGGTATGTT CAAGCAGCAT ACACTATGGA TATCCTCGGA 3360
AAGAGGAGGA AGATAATCTT CCTCGATGAA GCAGATAATA TAGAGCCCAG CGGAGCTAAG 3420

GAAATCGCAA AACTAATTGA TAAGGCCAAA AATCCAATAA TAATGGCTGC AAATAAGTAC 3480
TGGGAAGTTC CAAAAGAGAT CCGAGAAAAA GCTGAGCTAG TAGAGTACAA GAGGTTAACC 3540
CAGAGAGATG TAATGAATGC CTTAATAAGG ATCC 3574

SEQ ID NO: 84
SEQUENCE LENGTH: 33
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TACTTGTAAT ATTCTCATAT GATTGGCTCA ATA 33

SEQ ID NO: 85
SEQUENCE LENGTH: 35
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GATGAGTTCG TGTCCGTACA ACTGGCGTAA TCATG 35

SEQ ID NO: 86
SEQUENCE LENGTH: 25
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GGTTATCGAA ATCAGCCACA GCGCC 25

SEQ ID NO: 87
SEQUENCE LENGTH: 23
SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GCGTACCTTT GTCTCACGGG CAA 23

SEQ ID NO: 88
SEQUENCE LENGTH: 22
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GATAGCTGTC GTCATAGGAC TC 22

SEQ ID NO: 89
SEQUENCE LENGTH: 23
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
CTTAACCAGT GCGCTGAGTG ACT 23

SEQ ID NO: 90
SEQUENCE LENGTH: 28
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GACAATCTGG AATACGCCAC CTGACTTG 28

SEQ ID NO: 91

SEQUENCE LENGTH: 28
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TTGCCACTTC CGTCAACCAG GCTTATCA 28

SEQ ID NO: 92
SEQUENCE LENGTH: 29
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TGTCCGTCAG CTCATAACGG TACTTCACG 29

Claims

1. A thermostable DNA polymerase-associated factor capable of enhancing DNA synthesizing-activity of a DNA polymerase.

2. The DNA polymerase-associated factor according to claim 1, further possessing an activity of binding to a DNA polymerase.

3. The DNA polymerase-associated factor according to claim 2, which possesses an activity of binding to a DNA polymerase comprising a DNA polymerase-constituting protein having the amino acid sequence as shown in SEQ ID NO: 5 or 6 in Sequence Listing.

4. The DNA polymerase-associated factor according to any one of claims 1 to 3, comprising at least one of amino acid sequences selected from the group consisting of SEQ ID NOs: 1, 3, 19, 27, 34, 64, 70 and 80 in Sequence Listing, or an amino acid sequence resulting from substitution, deletion, addition or insertion of one or more amino acids in at least one of said amino acid sequences.

5. A gene encoding a DNA polymerase-associated factor, wherein the factor comprises at least one of amino acid sequences selected from the group consisting of SEQ ID NOs: 1, 3, 19, 27, 34, 64, 70 and 80 in Sequence Listing, or an amino acid sequence resulting from substitution, deletion, addition or insertion of one or more amino acids in at least one of said amino acid sequences, and possesses an activity of enhancing DNA synthesizing-activity of a DNA polymerase.

6. The gene according to claim 5, comprising a nucleotide sequence selected from the group consisting of SEQ ID NOs: 2, 4, 18, 26, 33, 63, 69 and 79, or a nucleotide sequence resulting from substitution, deletion, addition or insertion of one or more bases in said nucleotide sequence.

7. A gene capable of hybridizing to the gene of claim 5 or 6, and encoding a DNA polymerase-associated factor possessing an activity of enhancing DNA synthesizing-activity of a DNA polymerase.

8. A method for producing a DNA polymerase-associated factor, characterized in that the method comprises culturing a transformant harboring the gene of any one of claims 5 to 7, and collecting a thermostable DNA polymerase-associated factor capable of enhancing DNA synthesizing-activity of a DNA polymerase from the cultured medium.

9. A method of DNA synthesis by using a DNA polymerase, characterized in that DNA is synthesized in the presence of the DNA polymerase-associated factor of any one of claims 1 to 4.

10. The method of DNA synthesis according to claim 9, wherein DNA is synthesized in the presence of two or more kinds of DNA polymerase-associated factors.

11. The method of DNA synthesis according to claim 10, wherein DNA is synthesized in the presence of F7, PFU-RFC and PFU-RFCLS as a DNA polymerase-associated factor.

12. The method of DNA synthesis according to any one of claims 9 to 11, wherein said DNA polymerase is a thermostable DNA polymerase.

13. The method of DNA synthesis according to claim 12, wherein the synthesis is carried out by PCR method.

14. A kit usable for *in vitro* DNA synthesis, comprising the DNA polymerase-associated factor of any one of claims 1 to 4 and a DNA polymerase.

15. The kit according to claim 14, further comprising a reagent required for DNA synthesis.

16. The kit according to claim 14 or 15, comprising two or more kinds of DNA polymerase-associated factors.

17. The kit according to claim 16, comprising F7, PFU-RFC and PFU-RFCLS as a DNA polymerase-associated factor.

18. The kit according to any one of claims 14 to 17, comprising a thermostable DNA polymerase as a DNA polymerase.

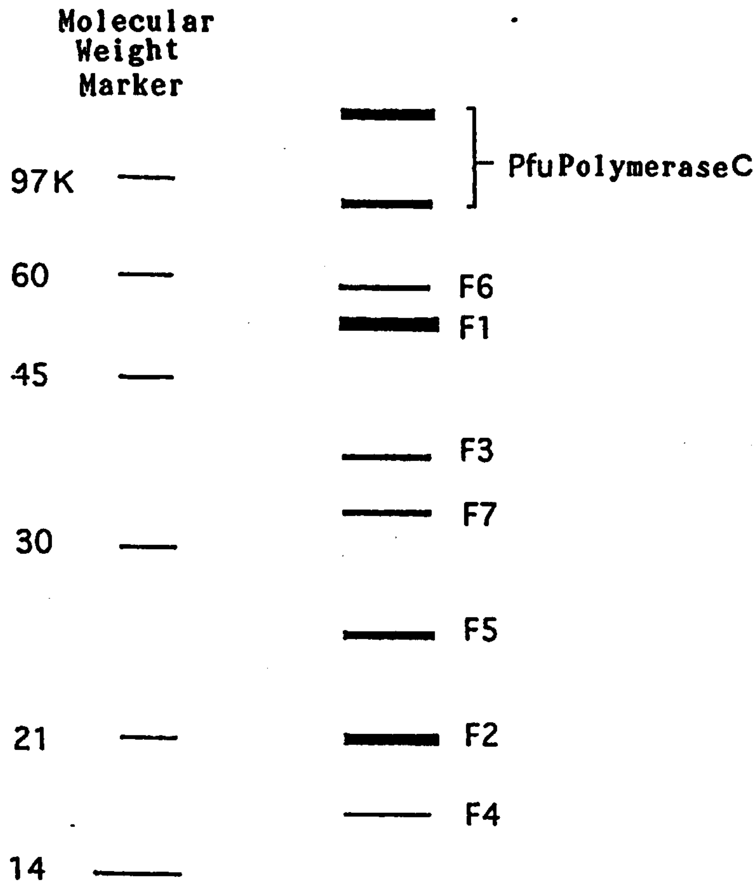
Molecular
Weight
Marker
97K
60
45
30
21
14
PfuPolymeraseC
F6
F1
F3
F7
F5
F2
F4

FIG. 1

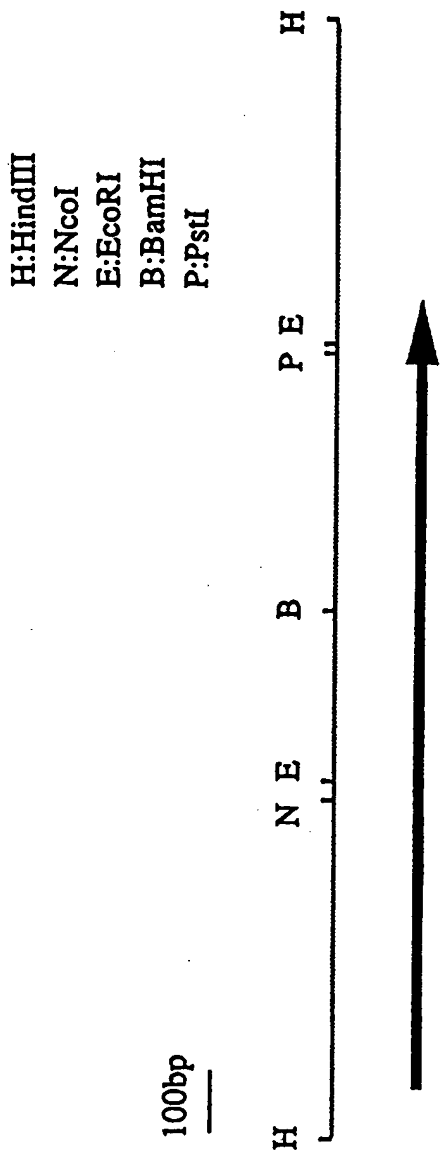

FIG. 2

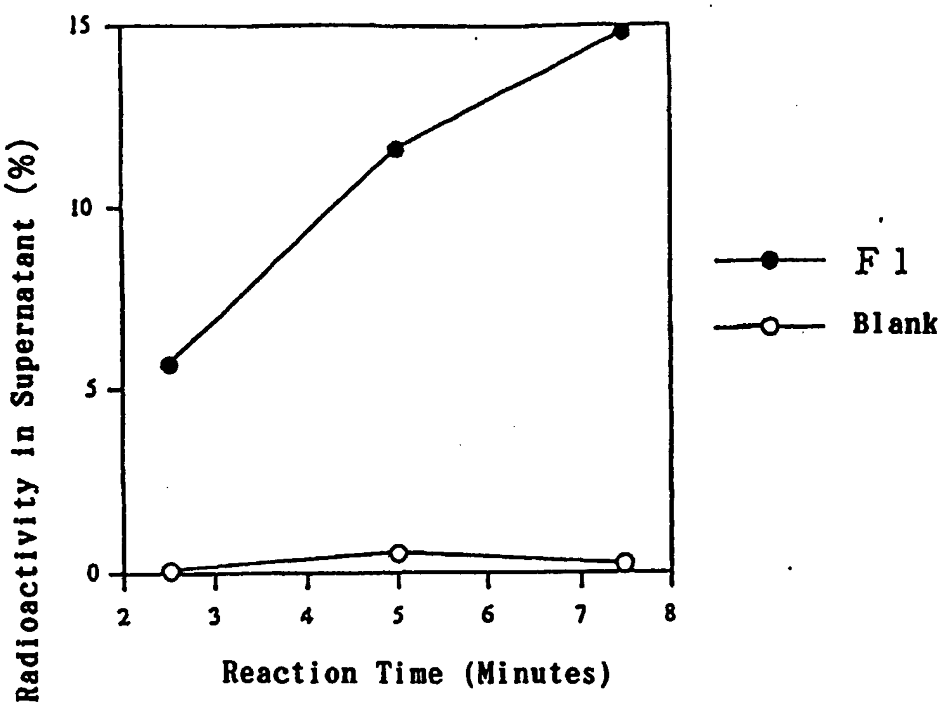
15
10
5
0
2
3
4
5
6
7
8
Radioactivity in Supernatant (%)
Reaction Time (Minutes)
F 1
Blank

FIG. 3

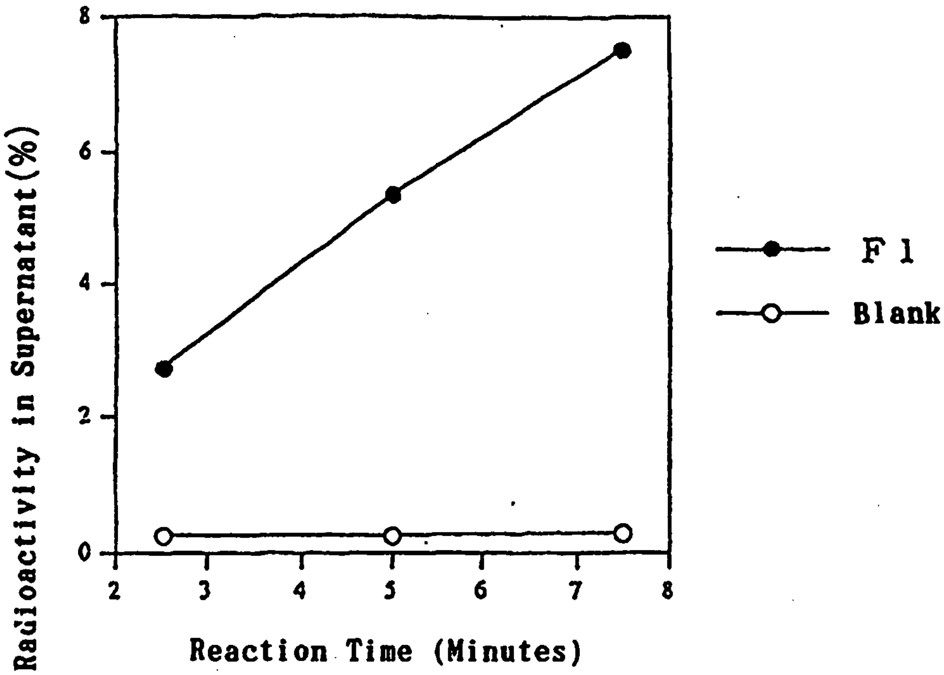
Radioactivity in Supernatant (%)
8
6
4
2
0
2
3
4
5
6
7
8
Reaction Time (Minutes)
F 1
Blank

FIG. 4

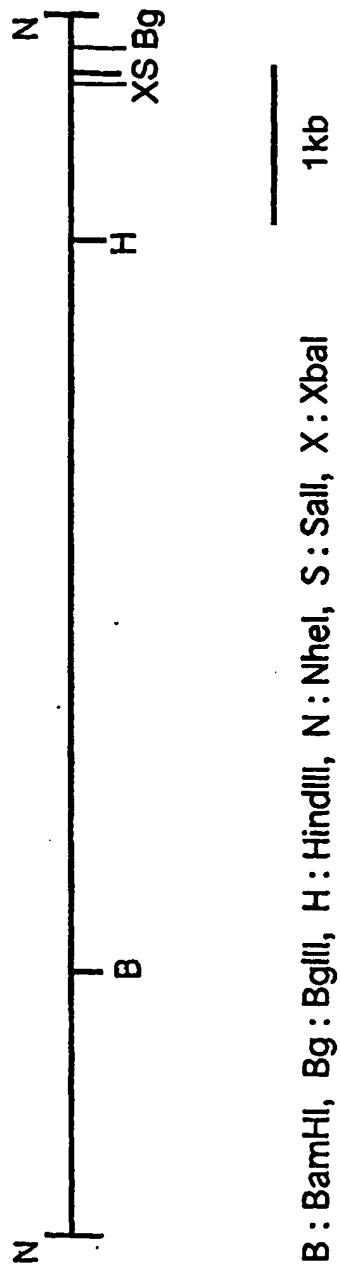

FIG. 5

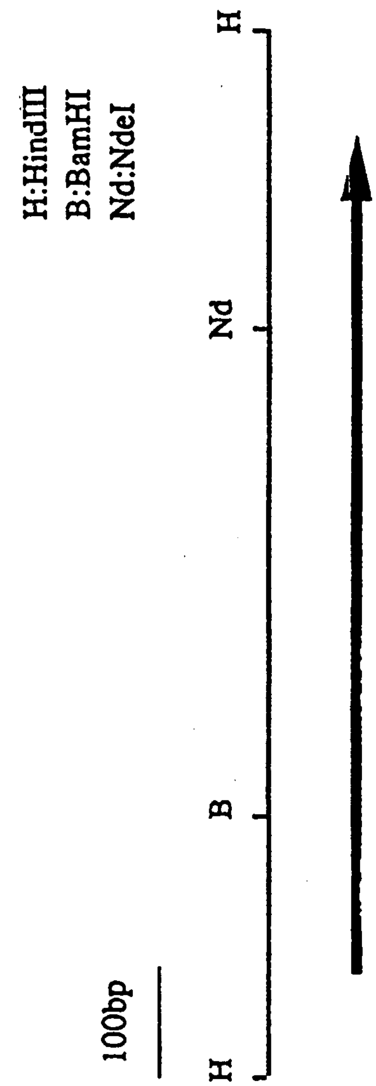

FIG. 6

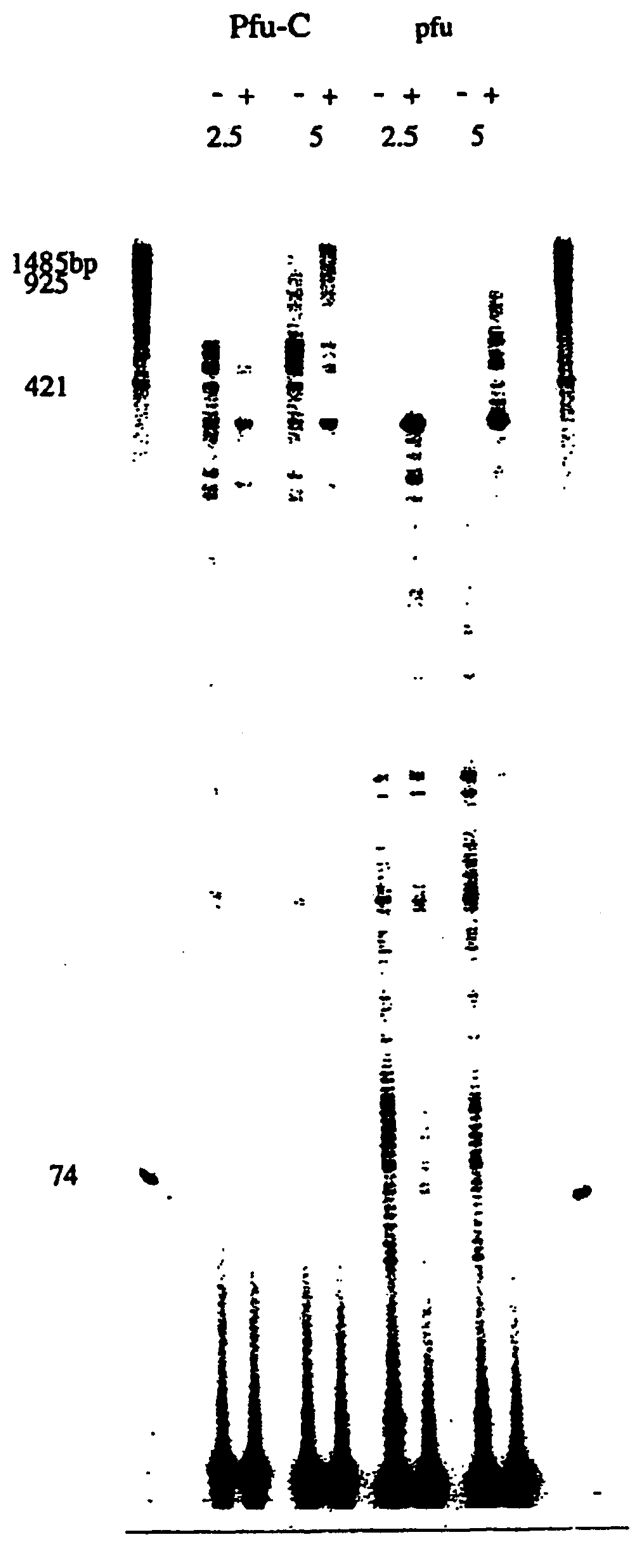

FIG. 7

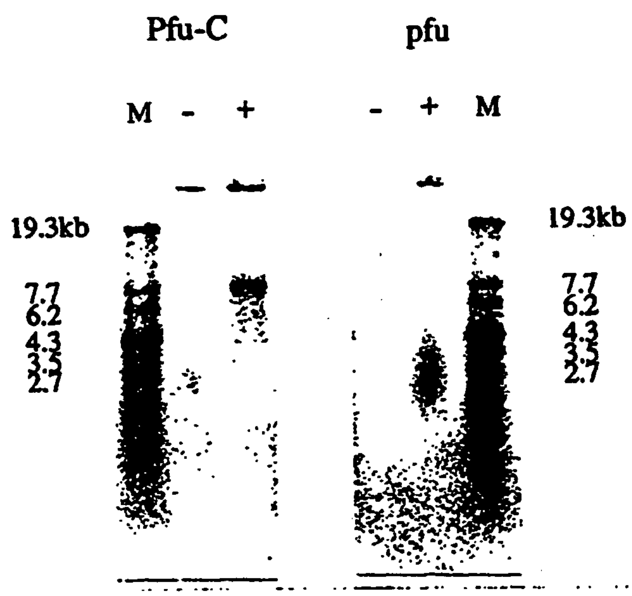

FIG. 8

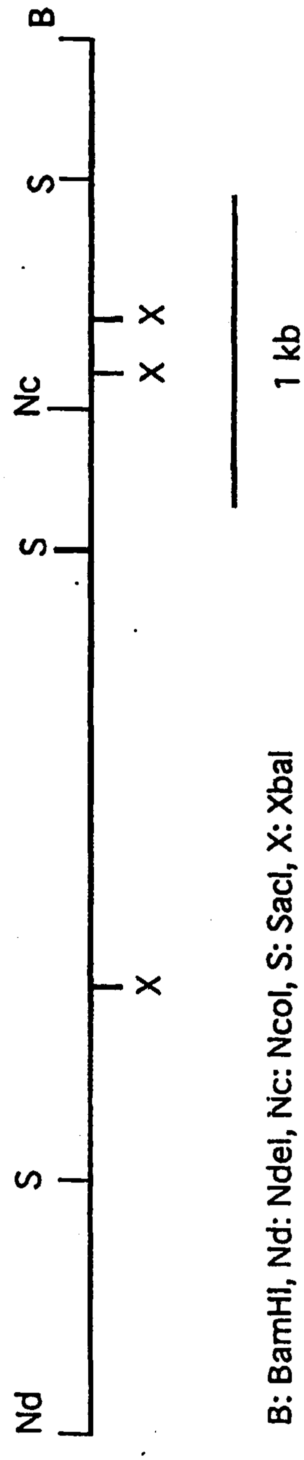

FIG. 9

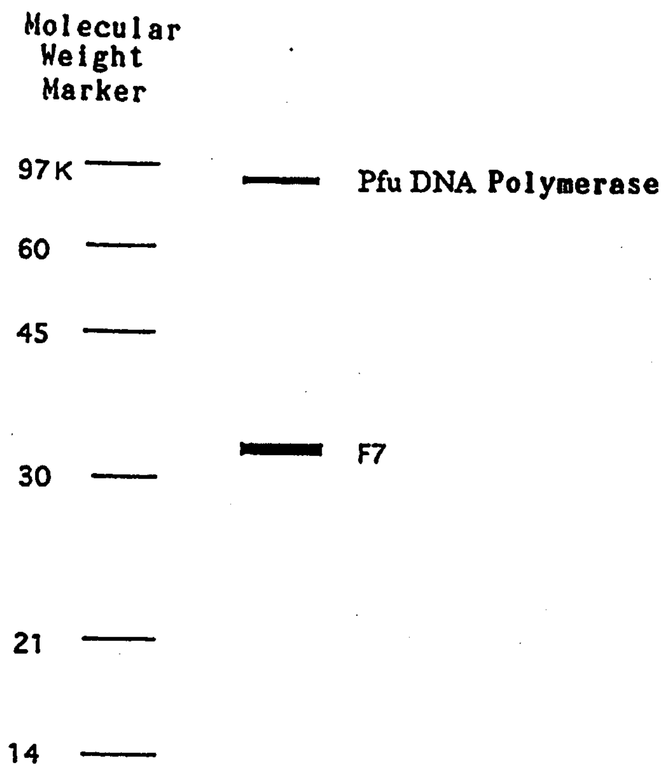
Molecular
Weight
Marker
97K
60
45
30
21
14
Pfu DNA Polymerase
F7

FIG. 10

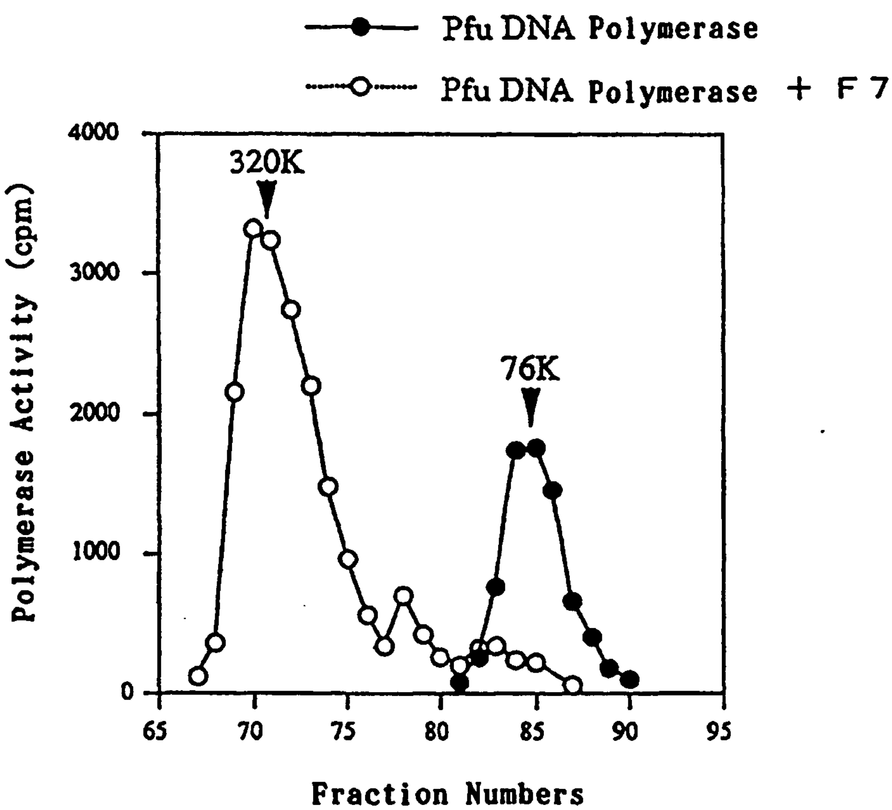
Pfu DNA Polymerase
Pfu DNA Polymerase + F 7
320K
76K
Polymerase Activity (cpm)
0
1000
2000
3000
4000
65
70
75
80
85
90
95
Fraction Numbers

FIG. 11

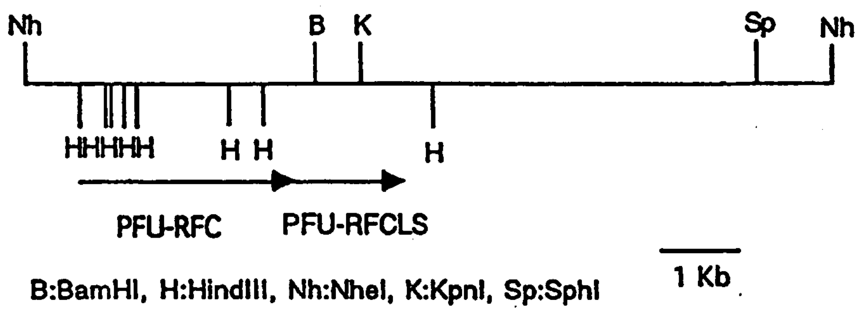
Nh
B
K
Sp
Nh
HHHHH
H H
H
PFU-RFC
PFU-RFCLS
B:BamHI, H:HindIII, Nh:NheI, K:KpnI, Sp:SphI
1 Kb

FIG. 12

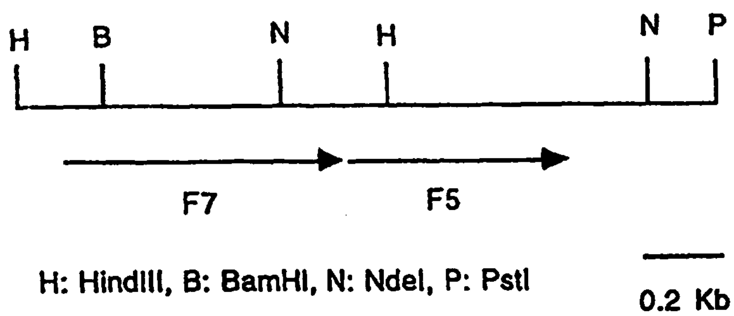

FIG. 13

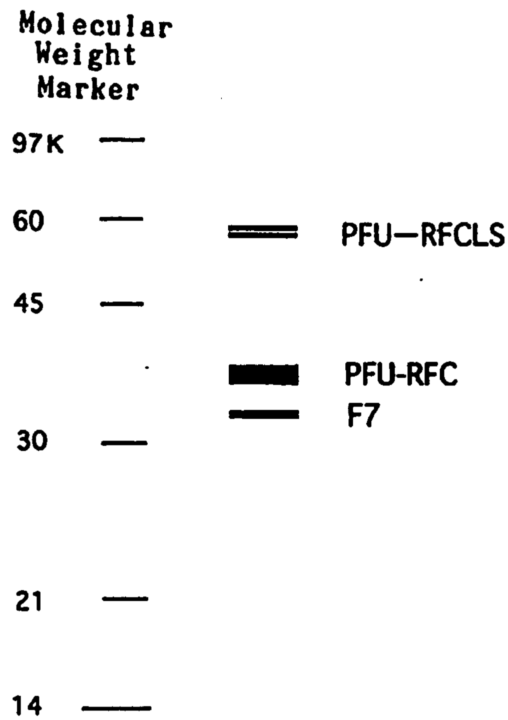
Molecular
Weight
Marker
97K
60
45
30
21
14
PFU—RFCLS
PFU-RFC
F7

FIG. 14

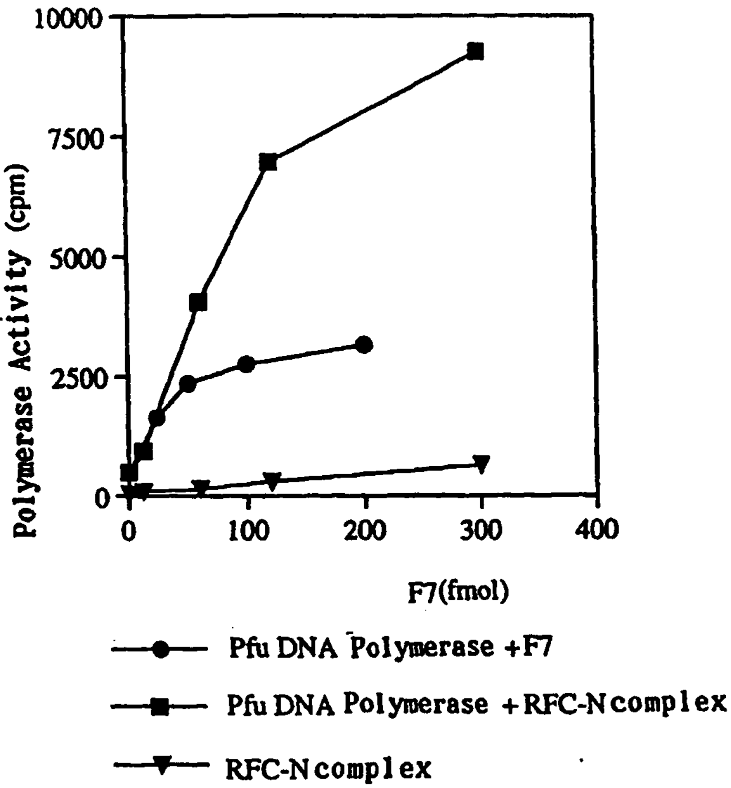
10000
7500
5000
2500
0
Polymerase Activity (cpm)
0
100
200
300
400
F7(fmol)
Pfu DNA Polymerase +F7
Pfu DNA Polymerase +RFC-Ncomplex
RFC-Ncomplex

FIG. 15

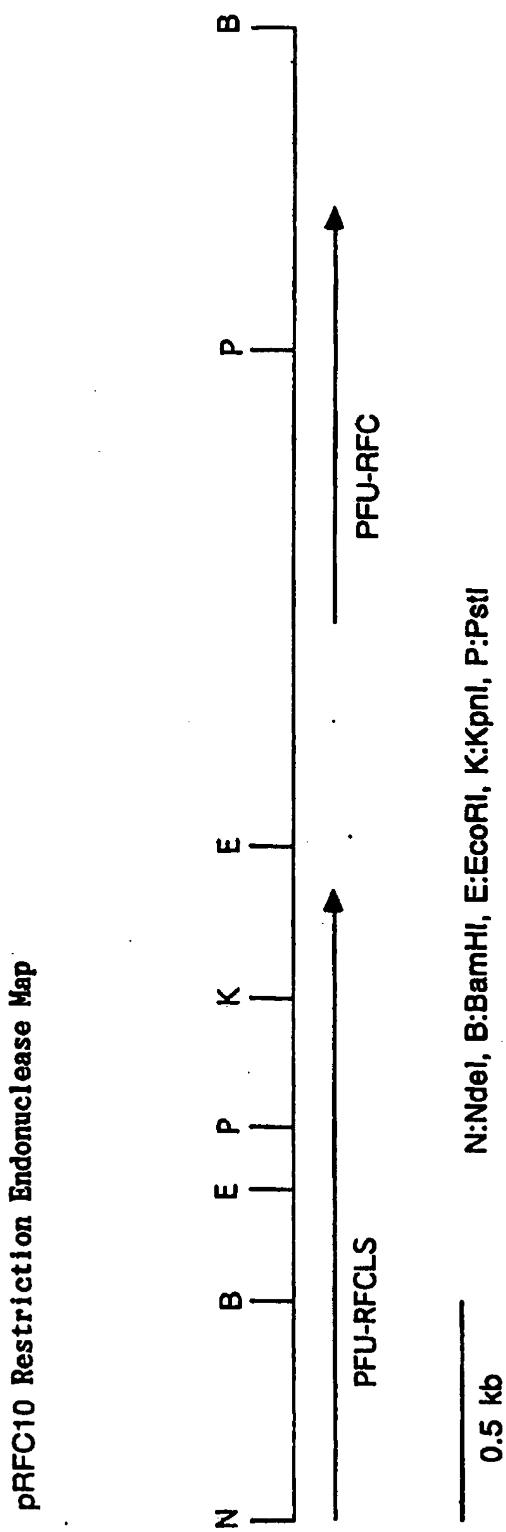
pRFC10 Restriction Endonuclease Map
N
B
E
P
K
E
P
B
PFU-RFCLS
PFU-RFC
N:NdeI, B:BamHI, E:EcoRI, K:KpnI, P:PstI
0.5 kb

FIG. 16

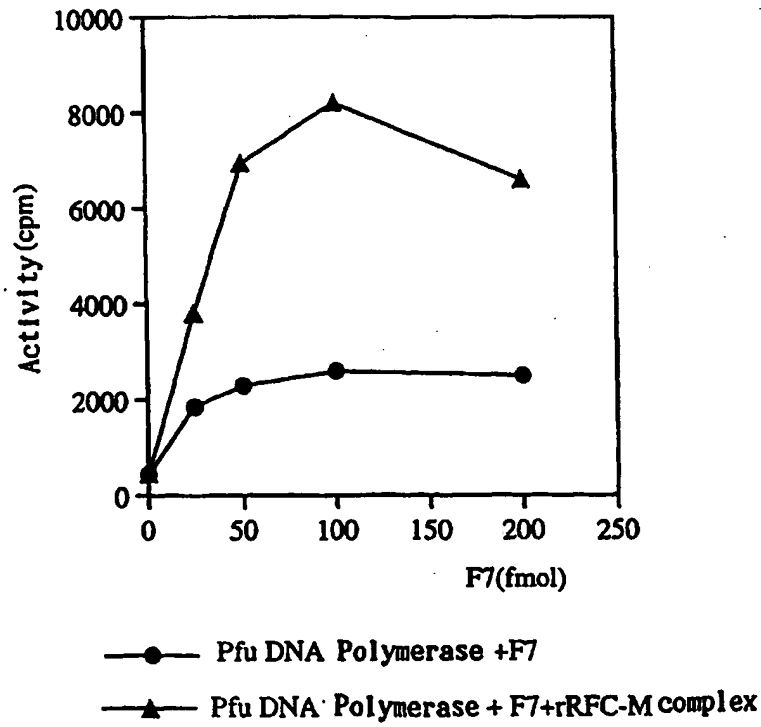
10000
8000
6000
4000
2000
0
Activity (cpm)
0
50
100
150
200
250
F7(fmol)
Pfu DNA Polymerase +F7
Pfu DNA Polymerase + F7+rRFC-M complex

FIG. 17

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP98/02845

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl⁶ C12N15/54, C12N9/12, C12N15/31, C07K14/195, C12P21/02 //
(C12N15/54, C12R1:01) (C12N9/12, C12R1:19) (C12N15/31, C12R1:01)
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
Int.Cl⁶ C12N15/54, C12N9/12, C12N15/31, C07K14/195, C12P21/02

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Genbank/EMBL/DDBJ/GeneSeq, WPI (DIALOG), BIOSIS (DIALOG)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX /PA	JP, 10-84954, A (The Institute of Physical and Chemical Research), 7 April, 1998 (07. 04. 98) & EP, 821058, A2	1, 9-10, 12-16, 18 /2-8, 11, 17
A	The Journal of Japanese Biochemical Society Vol. 68, No. 9 (1996) Hirosi Morioka "Structure and function of proliferating cell nuclear antigen (PCNA)" p.1542-1548	1-18

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 September, 1998 (25. 09. 98)	6 October, 1998 (06. 10. 98)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP98/02845

A. (Continuation) CLASSIFICATION OF SUBJECT MATTER

(C12P21/02, C12R1:19)

Form PCT/ISA/210 (extra sheet) (July 1992)